

Appendix H

835 ATACTCATGTGCAATCTCCTA 859

SULT 11

429436

XCUS

DEFINITION

AA429436 588 bp mRNA linear EST 16-OCT-1997
Z47212.1 Soares_total_fetus_Nb2HF8_9w Homo sapiens CDNA clone
IMAGE:773182.5 similar to WP:46H5.4 CE04590 RAT TRG GENE PRODUCT
mRNA sequence.

CESSION

ERSTION

EXORDS

ORUCE

ORGANISM

Homo sapiens

human.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 588)

Miller, L., Allen, M., Bowles, L., Dubuque, T., Gelsel, G., Jost, S.,

Kucaba, T., Lacy, M., Le, N., Lennon, G., Maira, M., Martin, J., Moore, B.,

Schellenberg, K., Steptoe, M., Tan, F., Theising, B., White, Y., Wylie

, T., Waterston, R. and Wilson, R.

Washu-Merck EST Project 1997

Unpublished (1997)

Contact: Wilson R.

Washington University School of Medicine

444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available royalty-free through LNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

Seq primer: -28ml3 rev2 ET from Amerham

High quality sequence stop: 495.

Location/Qualifiers

1. 588

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:773182"

/clone_lib="Soares_total_fetus_Nb2HF8_9w"

/dev_stage="8-9 weeks"

/lab_host="DH10B"

/note="Vector: pT7T3D-Pac (Pharmacia) with a modified

polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand CDNA

was prepared from mRNA obtained from pooled 8-9 week

(total) fetus material with a Not I - oligo(dt) primer [5'

TGTACCAATCTGAGTGGAGCGCGCTTAATTTTCTTTTCTTTT 3'].

Double-stranded CDNA was ligated to Eco RI adaptors

(Pharmacia), digested with Not I and cloned into the Not I

and Eco RI sites of the modified pT7T3 vector. Library

went through one round of normalization, and was

constructed by Bento Soares and M. Fatima Bonaldo."

202 a 99 c 124 g 163 t

BASE COUNT

ORIGIN

Query Match

Best Local Similarity 8.6%; Score 588; DB 9; Length 588;

Matches 588; Conservativity 100.0%; Pred. No. 5.2e-125;

Mismatches 0; Indels 0; Gaps 0;

QY 744 accagatgaggaagcaatagagagcttagtctctgatatacccaagaacattt 803

DB 1 ACCAGATGAGGAAGCAATAGAGAGCTTAGTCTCTGATATACCAGAACATTT 60

QY 804 tggtaaaagactctttaaataatgctatcactcaagtttgaattgaattgaacccat 863

DB 61 TGGTCAAAGACTCTCTTAATAATGCTATCACTCAAGTTGAATGAATGAACCCAT 120

QY 864 tttycaagtttgcttataatgatgcaaggaagaaagatttcagaaacttta 923

DB 121 TTTGCAAGTTTGGCTTTATATGATGTCAGGAAAGAAAGATTTCAGAAACTTTTA 180

QY 923 tttagaccttaattctgagcagatgaaggggttcttaagtcacacatgtaccacctgtgc 983

DB 181 TTTGACCTTAATTCAGAGATGAAAGGCTGTGTACGTCACATGTACCACTGCTGC 240

QY 984 cattaaccctggcgaagatcagcaatttttctatcactatccctcccaagatgttt 1043

DB 241 CATTACTACCTTGCGAAGATCAGCAATTTTCTATCATTATCCTTCCCAAGATGTTT 300

QY 1044 tctgtataaagctagaaaaaagtcctacagcaagagagacattggaagtgtgcagaacc 1103

DB 301 TCTGTATTAAGCTAGAAAAAGTCCCTACAGCAAGAGACATTGGAGAGTGTGCAAGACC 360

QY 1104 atatatatttccaagaagcagatgcccaccaagaataaagaaaaaactggagaaactgaa 1163

DB 361 ATATATGATTTTCAAGAGAGCAGATGCCACCAAGATAAGAAAACTGGAGAACTGAA 420

QY 1164 gattcaagcagatcagtttggccaaagacttgggaatatatcgcatgcttcttggac 1223

DB 421 GAGTCAAGCAGATCAGTTTGGCCAAAGACTTGGGAATATCGCATGCCCTTTGCTTGAC 480

QY 1224 tgcatacatltaataatgataattgttagcagtgctggaagtttggaaagagatctacaga 1283

DB 481 TGCATCCATTTTATGAATATTGTTAGCAGTGTCTGGAGATTGGAAAGAGATTCTACAGA 540

QY 1284 agtagaatcagtlactgagagaacgaaaggtcttggcagagagagag 1331

DB 541 AGTAGAATCAGTACTGAGAGACGAAAGGGTCTTGGTCAGAGAGAGAG 588

1510 bp mRNA linear HTC 03-JAN

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 16, 2002, 08:29:13 ; Search time 4715.69 Seconds
(without alignments)
19542.678 Million cell updates/sec

Title: US-09-737-246-1
Perfect score: 6828
Sequence: 1 gtcgcgcgtcgcgcagcagc.....ctcgtgcgcgcagcgcctc 6828

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*

1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1156.8	16.9	2710	11 AK016495	AK016495 Mus muscu
2	1044.4	15.3	1311	11 BC018599	BC018599 Mus muscu
3	678.8	9.9	898	10 BF309712	BF309712 601891854
4	678	9.9	1873	11 BC018075	BC018075 Homo sapi
5	678	9.9	1873	11 BC018076	BC018076 Homo sapi
6	662	9.7	2694	11 AK004816	AK004816 Mus muscu
7	651.4	9.5	795	10 BI088758	BI088758 602851547
8	646	9.5	1051	11 AK014226	AK014226 Mus muscu
9	610.4	8.9	931	10 BG247108	BG247108 602359776
10	599.4	8.8	958	10 BE891826	BE891826 601434272
11	588	8.6	588	9 AA429436	AA429436 zw47a12.r
12	580.6	8.5	1510	11 BC020473	BC020473 Homo sapi
13	572	8.4	869	10 BI689126	BI689126 603315930
14	565.6	8.3	796	9 AJ397707	AJ397707 AJ397707
15	561.4	8.2	745	10 BG872239	BG872239 602790848
16	549.8	8.1	575	10 BM310731	BM310731 ig47g06.y
17	543.2	8.0	636	11 AK013336	AK013336 Mus muscu

18	543	8.0	678	10 BI101016	BI101016 602886468
19	542.6	7.9	606	10 BF036757	BF036757 601459988
20	541	7.9	645	10 BF001267	BF001267 799610.x
21	539.6	7.9	884	10 BG677746	BG677746 602625421
22	539.2	7.9	1223	10 BG922088	BG922088 602386380
23	535.2	7.8	540	10 BE698863	BE698863 RC4-NN002
24	527.8	7.7	928	10 BG205542	BG205542 RST24921
25	523.8	7.7	935	10 BF159615	BF159615 601767221
26	520.8	7.6	636	9 AW919177	AW919177 EST350481
27	511	7.5	620	10 BG927424	BG927424 HNC53-1-E
28	492.2	7.2	662	10 BI732553	BI732553 603355483
29	489.4	7.2	547	10 BM310431	BM310431 ig47g06.x
30	482	7.1	536	9 AW812481	AW812481 CML-ST018
31	475.4	7.0	640	9 AA174111	AA174111 zp19h01.s
32	472.2	6.9	737	10 BF141712	BF141712 601790772
33	471	6.9	648	10 BF120001	BF120001 601758086
34	467.2	6.8	492	9 AA933066	AA933066 om85e01.s
35	467.2	6.8	503	10 BF436753	BF436753 7p09h06.x
36	467.2	6.8	529	10 BF515907	BF515907 UI-H-BW1-
37	465.4	6.8	488	9 AW057724	AW057724 wx02g04.x
38	462.8	6.8	506	9 AA777760	AA777760 z195g04.s
39	459.8	6.7	567	9 AA275395	AA275395 va92f11.r
40	456.6	6.7	595	10 BG080811	BG080811 H3057G07-
41	455.2	6.7	514	10 H98193	H98193 yx08d10.s1
42	450.6	6.6	472	9 BE219721	BE219721 hv61f05.x
43	449.2	6.6	507	9 AA310291	AA310291 EST181118
44	447.4	6.6	537	10 BI491205	BI491205 dF05d04.w
45	445.4	6.5	477	9 AW024109	AW024109 wu62e08.x

ALIGNMENTS

RESULT 1	AK016495	2710 bp	mRNA	linear	HTC 19-JAN-2002
LOCUS	Mus musculus adult male testis cDNA, RIKEN full-length enriched library, clone:4931431C02:homolog to KIAA1395 PROTEIN (FRAGMENT), full insert sequence.				
DEFINITION	AK016495				
ACCESSION	AK016495.1 GI:12855259				
VERSION	AK016495				
KEYWORDS	HTC; CAP trapper.				
SOURCE	Mus musculus (strain:C57BL/6J) adult male testis cDNA to mRNA, clone_lib:RIKEN full-length enriched mouse cDNA library clone:4931431C02.				
ORGANISM	Mus musculus				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.				
AUTHORS	1 (sites)				
TITLE	Carninci, P. and Hayashizaki, Y.				
JOURNAL	High-efficiency full-length cDNA cloning				
MEDLINE	Meth. Enzymol. 303, 19-44 (1999)				
PUBMED	99279253				
REFERENCE	10349636				
AUTHORS	2 (sites)				
TITLE	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Konno, H., Akiyama, J., Nishi, K., Kitsunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.				
JOURNAL	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer				
TITLE	Genome Res. 10 (11), 1757-1771 (2000)				

MEDLINE	20530913
PUBMED	11076861
REFERENCE	4 (sites)
AUTHORS	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
TITLE	Functional annotation of a full-length mouse cDNA collection
JOURNAL	Nature 409, 685-690 (2001)
REFERENCE	5 (bases 1 to 2710)
AUTHORS	Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A., Arakawa,T., Baldarelli,R., Bono,H., Brownstein,M., Bult,C., Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hill,D., Hiramoto,K., Hiraoka,T., Horii,F., Hume,D., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Quackenbush,J., Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Schriml,L., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yamanaka,I., Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
TITLE	Direct Submission
JOURNAL	Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gs.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
COMMENT	Please visit our web site (http://genome.gsc.riken.go.jp/) for further details. cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGAGAGAGAGATCCAGAGAGCTCTTTTCTTTTCTTTTCTTTVN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 100.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCGAGTTAATTAATTAATTCACCCCCCCCC 3']. cDNA was cleaved with BamHI and XhoI. cDNA of size comprised longer than 7 kb was selected before cloning. Vector: a modified Bluescript KS(+) after bulk excision from lambda FLC I. Cloning sites, 5' end: SalI; 3' end: BamHI. Host: DH10B.
FEATURES	location/Qualifiers
source	1. 2710
	/organism="Mus musculus"
	/strain="C57BL/6J"
	/db_xref="MGD:MGI:1897080"
	/db_xref="taxon:10090"
	/clone="4931431C02"
	/sex="male"
	/tissue_type="testis"
	/clone_lib="RIKEN full-length enriched mouse cDNA library"
	/dev_stage="adult"
	<1. 2536
	/note="data source:SPNR, source key:Q9P2F2, evidence:ISS homolog to KIAA1395 PROTEIN (FRAGMENT)"
	putative"
	/codon_start=2
	/protein_id="BAB30271.1"
	/db_xref="GI:12855260"
	/translation="RGHRKHOPICGHHRWGSPGWSLHQHPCTFNSSSLRLPLSA
	ESSRTLVCVLWLKNAEPTLLQWAAIDLALPOLGRLLDLVLCIAFEYKGKKAER
	INSLTFKSLDMKARLEALITGARQEMVRKRSRSPFGQENYVWRKKSATHWRQT
	SDRVDTKDEMEHEALVDGNLATEASLVLDLTLETIVQTVMLSEARESILSAVLKVL
	YSLGSAQSAFLQHGALQARALVSKPELLFEEDTELCAIDLRLRHGSRISTIRM
	HASASLYIMRONFEIGHNFAVRYKMLVETMSLSLVGTTQNFSEHRLRSKLTITYAE
	EDIGLRDSTFAEQVODLMFNHMLITDTVKMKEHQEDPEMLMDIMYRIARGYGSPDL
CDS	

RY	4694	caatacaggtcacaccccaagtcctcccttaccactaatgagcgaaaactttagatctg	4753
QY	4694	caagagcagtgctgctatattatgcctcaggtctctccgacactgtagcagtagcatcgta	4693
Db	759	CGGAGCTTTGTGCGGACCTGTGCTGAGACTTCTGCGACACTGTGGCAGCCGCATCAGCA	818
QY	4574	gttttgctacacagagagccttggttccaagtctcctgaactcttatttgaagaaga	4633
Db	699	GCCTGGCCACCCAGCGGGCCTGGTCTCCAAGTTTCCGGAGCTGCTTTTCGAGGAAGACA	758
QY	4514	tgctaaaagtgtctactacacagcatgycctgttaaccaaaagtgagtttatactacaacact	4573
Db	639	TGCTGAAGTGTCTCTTACAGTCTTGGGAGCGCCCAAGAGTGCCCTGTCTTCGACGATG	698
QY	4454	tagagatgtgttctcagaccggttctgttaacggaatccaagaagagcatctctgtgtgag	4513
Db	579	TGGAGACCATCGTGCAGACAGTGTGCTGTCCGAGGCCCGTGAGAGCATCTCGAGTGGCG	638
QY	4394	acgaagcactgatgtatgtaaacctgtgtcacaagaagcaacctaatactttagatacat	4453
Db	519	ATGAAGCTCTGGTGGACGGAACCTGGCGACAGAGCCAGCCTGGTGTCTTGACACGC	578
QY	4334	atatgactcactgycgttcaaaaacacagagaagcttgacaacaatcaagagcagagattgaac	4393
Db	459	GTGCCACACACTGGAGACAGACCTCAGACCCTGTGTGACCAAGACCAGATGAATGGAAC	518
QY	4274	tcgagagaagcccatctgtgaagtgcccttgaagtcaagaaatltgaagltggaagaag	4333
Db	417	-----GAAGTCCATTTTGGAACCAAGAGAAATGTTCCTGGCGGAAGA	458
QY	4214	aagctattcttgggagcatagtggtccagycgaagaatggtacgycgaagccgagagcagc	4273
Db	360	AAGCCATCTGGCGCACCATTTGGGGCGCGGACGAGATGGTGGCGGAGTGGGAGA---	416
QY	4154	ttgaacgaatgaatagcttgacctttaagaatacaaaagacatgtagagcaaaagcttgaag	4213
Db	300	TTGAGCGGATCAACAGCCTCACGTTCAAGAAGTCTCTGGACATGAALAGCCGGCTGGAGG	359
QY	4094	ggctattagatctgttattctctgtgtgtcttgtttagtataaagggaaaagtggt	4153
Db	240	GACTCCTCGACTTACTTACTCTGTCTGTGGCTGCTTCGAGTACAAGGGGAAGAAAGCCT	299
QY	4034	atgcagatbaaacagttctacagaagltggtttacagatctctcagttctgacgttaaac	4093
Db	180	ATGCAGAGCCACGCTCCTGCAGCGCTGGGCTGTGACTGCGCCCTGCCCAAGCTGGGGC	239
QY	3974	ctaccttttcagcagaatcaagtggaagccttttgatctgtctacttgggttctcaaaa	4033
Db	120	TTCTCTCTCTGTGTAATCCAGCCGGACCTTGCTGTGTGTGTCTCTGTGGGTCTGAAGA	179
QY	3854	aaagtgaagcggaagtatgataagccagaccgttgccatgccaatcgcaggagacatcgy	3913
Db	1	AAGGGACATCGGAAGCACCATCATCAACCCATCTGTGGCCATGGCCATCGCTGGGGGTCCCC	60
QY	3914	tcctcctaactacaagcgctgtgcagtttccctcctaagtcacaagagtgycagycacaaca	3973
Db	61	TGGCTCCTGGCTCTCGCACCAAGCATTTCCCAAGGAGACCTTTCAAC-AGCAGCTTCGCTCAGGC	119
QY	3974	ctaccttttcagcagaatcaagtggaagccttttgatctgtctacttgggttctcaaaa	4033
Db	120	TTCTCTCTCTGTGTAATCCAGCCGGACCTTGCTGTGTGTGTCTCTGTGGGTCTGAAGA	179
QY	4034	atgcagatbaaacagttctacagaagltggtttacagatctctcagttctgacgttaaac	4093
Db	180	ATGCAGAGCCACGCTCCTGCAGCGCTGGGCTGTGACTGCGCCCTGCCCAAGCTGGGGC	239
QY	4094	ggctattagatctgttattctctgtgtgtcttgtttagtataaagggaaaagtggt	4153
Db	240	GACTCCTCGACTTACTTACTCTGTCTGTGGCTGCTTCGAGTACAAGGGGAAGAAAGCCT	299
QY	4154	ttgaacgaatgaatagcttgacctttaagaatacaaaagacatgtagagcaaaagcttgaag	4213
Db	300	TTGAGCGGATCAACAGCCTCACGTTCAAGAAGTCTCTGGACATGAALAGCCGGCTGGAGG	359
QY	4214	aagctattcttgggagcatagtggtccagycgaagaatggtacgycgaagccgagagcagc	4273
Db	360	AAGCCATCTGGCGCACCATTTGGGGCGCGGACGAGATGGTGGCGGAGTGGGAGA---	416
QY	4274	tcgagagaagcccatctgtgaagtgcccttgaagtcaagaaatltgaagltggaagaag	4333
Db	417	-----GAAGTCCATTTTGGAACCAAGAGAAATGTTCCTGGCGGAAGA	458
QY	4334	atatgactcactgycgttcaaaaacacagagaagcttgacaacaatcaagagcagagattgaac	4393
Db	459	GTGCCACACACTGGAGACAGACCTCAGACCCTGTGTGACCAAGACCAGATGAATGGAAC	518
QY	4394	acgaagcactgatgtatgtaaacctgtgtcacaagaagcaacctaatactttagatacat	4453
Db	519	ATGAAGCTCTGGTGGACGGAACCTGGCGACAGAGCCAGCCTGGTGTCTTGACACGC	578
QY	4454	tagagatgtgttctcagaccggttctgttaacggaatccaagaagagcatctctgtgtgag	4513
Db	579	TGGAGACCATCGTGCAGACAGTGTGCTGTCCGAGGCCCGTGAGAGCATCTCGAGTGGCG	638
QY	4514	tgctaaaagtgtctactacacagcatgycctgttaaccaaaagtgagtttatactacaacact	4573
Db	639	TGCTGAAGTGTCTCTTACAGTCTTGGGAGCGCCCAAGAGTGCCCTGTCTTCGACGATG	698
QY	4574	gttttgctacacagagagccttggttccaagtctcctgaactcttatttgaagaaga	4633
Db	699	GCCTGGCCACCCAGCGGGCCTGGTCTCCAAGTTTCCGGAGCTGCTTTTCGAGGAAGACA	758
QY	4634	cagagcagtgctgctatattatgcctcaggtctctccgacactgtagcagtagcatcgta	4693
Db	759	CGGAGCTTTGTGCGGACCTGTGCTGAGACTTCTGCGACACTGTGGCAGCCGCATCAGCA	818
QY	4694	caatacaggtcacaccccaagtcctcccttaccactaatgagcgaaaactttagatctg	4753

Db 819 CCATCCGCATGCACGCGCCGCTCCCTCTACTGCTTATGCGCCAGAACTTCGAGATTG 878

QY 4754 ggaataactltgccaaggltaaatgcagtiaccaatgtcacatactcctctgtgtgca 4813

Db 879 GCCATAACTTGGCCGCTGTGAAGATGCTGGTGACCATGTCTGTGCTCCCTGTGGGGA 938

QY 4814 catctcgaattttaagaagaatttctaagagttctctaagactatataatgacatg 4873

Db 939 CAACTCAGAACTTTAGTGAAGAGACATTTGAGAAGTCCCTCAAGACCATCCTGAACCTACG 998

QY 4874 ctgaagaagatctggaattgagggaacaacacattccttgatcaggtccagatctgttt 4933

Db 999 CAGAGGAGGACATAGGGCTGAGGGGACAGCACCTTCCTGAGCAGAGTCCAGGACCTCATGT 1058

QY 4934 tcaatclocatatgattcttctgatactgtgaaataagaagaacacaggagatcctg 4993

Db 1059 TCAACCTGCACATGATCCTGACAGACACGCTGAAGATGAAGGAACACAGGAGACCCCTG 1118

QY 4994 aaatgtgatgtatctaagtacagaatgtccaagggttaccagacctccaga---gc 5050

Db 1119 AGATGCTCATGGACCTCATGTACAGAAATGCTCGGGGATACCAAGGCTCPCAGACCTGC 1178

QY 5051 gattgacctgtgtgcagaacatgagcaagcactcagaacgaagcaatcatgtctgaag 5110

Db 1179 GGCTGACATGGCTGCAGAACATGGCTGGCAAAACATGCAGAGCTGGGCAATCATGCAGAGG 1238

QY 5111 ctgcacagtgctagtccactcagcagcactgtgtctgaataatttgacatgtctgagg 5170

Db 1239 CTGCCAGTGCATGGTCCACGCCGCTGCCCTGGTGGCGAAATACCTTGCCCTCCTCGAGG 1298

QY 5171 accggaataatcttcctgtgagatgtgtaacattcagaataattcataatgttttag 5230

Db 1299 ACAGTCCGACACTGCGCTGTGGGCTGTGTCTCCATCCAGAAATGTCTCATCCAACGCTGG 1358

QY 5231 aagaatctgcgtctcagatgtgtatctccagatgaagaagtatctgtctgaa 5290

Db 1359 AGGAGTCTGCCATCTCTGATGACATCCTGTGCGCCAGATGAGGAGGGCTTCTGCTTAGGA 1418

QY 5291 aatacttactgagtcagactgtgggattactgtgaacaagcagctgtcctctcta 5350

Db 1419 AGAACTTCACAGAACTGGGTCTGGTGGGGCTGCTGGAGCAGGGCGGCTACTTCACCA 1478

QY 5351 tggctgcatgtatgaagcagttaatgaagtttacaagtacttattcctattcaatgaag 5410

Db 1479 TGGGTGCTCTGTATGAAGCGGTGAACGAAGTCTACAAAAAACCCTTATCCCCCATCTTGAAG 1538

QY 5411 ctaatcgagatgcaagaactataccacaattcatgtaaaccttcaagaagcattcaagca 5470

Db 1539 CCCACAGAGACTACAAGAAGCTGGCTGGCTGCACGGGAAACTGCAGGAGGCGCTTCACCA 1598

QY 5471 aaatgttcalcagagtagtgcgtgggagcgatgttggcacctatttcgtgtgtt 5530

Db 1599 AGATTATGCACACAGAGCTCTGGCTGGGAGCGGTGTATTGGGACATATTTCGAGTGGGCT 1658

QY 5531 ttataggaaccaagtgcgggattggatgaacaagaattgttacaagagcgctgcaa 5590

Db 1659 TCTATGGCACACGATTTGGTGACCTGGATGAACAAGAGTTGTGTACAAGAAGACCGTCAA 1718

QY 5591 taaccaacttgcaagatatctcacaagattgagggattttaagggaagaagattggag 5650

Db 1719 TCACGAGAGCTTGACAGATCTCACACCGGCTGGAGAGTCTTATACGGAAGGTTGCGGG 1778

QY 5651 aggatgtggtgaagtaatacaaaagcttaatcctgttagacaagtgtlaaatagatccta 5710

Db 1779 ATGATGTGTAGAGATCATCAAGATTTCTAACCAGTGGACAGTCCCAAGCTGGACCCAC 1838

QY 5711 acaagcatatatcagattacatagtgtgagcataacttgaacacatatgagatgaag 5770

Db 1839 AGAAGGCGTACATACAGATAACCTATGTGAGGCCACATTTGCACACTTATGAGCTCAAGG 1898

QY 5771 acagaatcaactattcgacaaaaatacaatcttcgtcgatcatgtactgtacacct 5830

Db 1899 ATCGGTGACCTACTTGCATCGGAACATATGGGCTGGGGCCTTCCCTCTTGACACACCT 1958

QY 5831 ttacttagatgcccgtgcccataggygaactlcatgaaacattcaaaagaagacattc 5890

Db 1959 TCACACCAGATGGACGTGGCGCAGGAGAGTTGGCCGACACACAAACGCAACGCTGC 2018

QY 5891 tgactagtlcatgaccttccattatataaacaagaagtcaatlytcaactcataaagaag 5950

Db 2019 TGAGCAGGAGCATGCTTTCCTTACATCAAGACACCGCATCCGAGTGTGCCACCGTAGG 2078

QY 5951 agatcattcaacaccaatltgaagltgtatltgaagcattgcagaanaagaacagagat 6010

Db 2079 AGACAGTCTGACACACAGTGAAGTGGCCATTGAGGACATGCAGAAGAAGACCGGAGC 2138

QY 6011 tggcattgcaacacatcagatcccgagaccccaaatgcttcagatgttactcagg 6070

Db 2139 TGGCCTTGCCACCGAGCAGACCTCCAGATGCCAAGATGCTGCAGATGTTCTCCAGG 2198

QY 6071 gatctgtaggcaccacagtgaaatcaaggcgcttggaaagtggccagtttctgtctg 6130

Db 2199 GTTCTGTGGGACCCACTGTGAACCAAGGTTCCCTTGAAGTGGCCAGGTGTTTGTGACG 2258

QY 6131 aaatacctagtgacccaaagctcttcagacatcataataactgcactctgtcttaag 6190

Db 2259 AGATCCAGAGATCCCAAGCTCTTCGACATCACAAACAACTCCGGCTCTGCTTCAAGG 2318

QY 6191 atttactaaaagltgtgaagatgcttaagaanaaaataagacttaattggtcggttc 6250

Db 2319 ATTTCTGCAAAAGTGCAGAGATGCAGTGAAGAAAGAACAAAGGCCCTGATTGGCCAGACC 2378

QY 6251 aaaagagttcaaaagggaattg---ggaaactatcttcgcttaagagggcctacagc 6308

Db 2379 AGAAGAGTACCAACCGGAGCTGAGGCTCACTATAGCCGCTGGCGGAGGCTCTGCAGC 2438

QY 6309 cctagatcacagaagtc 6326

Db 2439 CTCTGCTTACCCAACGTC 2456

RESULT 2

BC018599 1311 bp mRNA linear HTC 03-DEC-2001

LOCUS

DEFINITION Mus musculus, clone IMAGE:3596745, mRNA.

ACCESSION BC018599

VERSION BC018599.1 GI:17235789

KEYWORDS HTC.

SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 1311)

AUTHORS Strausberg,R.

TITLE Direct Submission

JOURNAL Submitted (03-DEC-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK

COMMENT NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgabs-r@mail.nih.gov

Tissue Procurement: Jeffrey Green M.D.

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center

Center code: BCM-HGSC

Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>

Contact: villalobcm.tmc.edu

Villalón, D.R., Luna, R.A., Hale, S.M., Huilyk, S., Lu, X., Garcia, A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W., Muzny, D.M., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 16 Row: e Column: 2
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis

This clone has the following problem: no 5' EST match.

location/Qualifiers

source
1.1311
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:3596745"
/tissue_type="Mammary tumor, C3(1)-Tag model, Infiltrating
ductal carcinoma, 5 month old virgin mouse."
/clone_lib="NCL CGAP_Mam6"
/lab_host="DH10B"
/note="Vector: PCMV-SPORT6"

BASE COUNT 363 a 309 c 333 g 306 t
ORIGIN

Query Match 15.3%; Score 1044.4; DB 11; Length 1311;
Best Local Similarity 87.9%; Pred. No. 1.9e-230;
Matches 1151; Conservative 0; Mismatches 156; Indels 3; Gaps 1;

QY 3832 tgtatagccactgatgatattgaagaagtgaagcggaagtatgataagccagaccgtgccc 3891
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 TGTATAGCCCTTGATGACTATGACAGTGAAGTGGAGCATGATAAGCCAGACGCTTGCC 60
QY 3892 atgccaatcgcaaggacatcgtgtccctcaactaacaaggcctgycagtttccctccacg 3951
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 ATGGCAATTGCAAGAACATCCGTCCCTCAGCTGACACAGACCTGGCAGTTTCTCTCACG 120
QY 3952 tcaacgagtggcaggaacaacactactcttccagcagaatcaagtgcgaagccttttgatc 4011
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 121 TCAACGAGTGGCCGGCAGCAGCACCACTTCTCAGCAGAGTCAAGCAGAGCCCTTTTGATC 180
QY 4012 tgtctactttgggttctcctcaaaaatgcagatgaacaagttctacagaagtgtttacagat 4071
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 181 TGTCTGCTTTGGGTTCTCAAGAAATGCAGATGAAGCGGTTCTGCAGAAATGGTTACAGAT 240
QY 4072 ctctcagctcttgcaactaaccggcctattagatctgtcttatctctgtgtcttctt 4131
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 241 CTCTCAGTCTCTCAGCTGAACCGGTTGTTAGATCTGCTTTACCTGTGTATCTTGCTTT 300
QY 4132 gaatataaagggaaaaagtgtttgaacgaatgaatagcttgacctttaagaatcaaaa 4191
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 301 GAGTACAAAGGAAAAAGGTATTGAAAGAAATGAATGATTAACTTTAAGAAATCAAAA 360
QY 4192 gacatgagagcaaaagcttgaagaagctattcttgggagcataggtgccaggaagaatg 4251
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 361 GACATGAGAGCCAAAGCTTGAGAGAACCATTTCTGGGAGCATCGGTGCCAGGCAAGAAATG 420
QY 4252 gtacggcggaagccgaggaacgtctgagagaagcccatctggaagtgcctttggaagtcaa 4311
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 421 GTACGGCGGAGCCGAGGGCAGCTCGAGAGGAGGCCCATCTGGAAGCGCCTTTGGAGGCCAA 480
QY 4312 gaaaatttgaggtggaagaagatatgactcactggtcgctcaaaaacacagagaagcttgac 4371
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 481 GAAAAAAGCTGGCGGTGAGAGAAAGACATGACTGCGCTCAGAACTCAGAGAAACTCGAT 540
QY 4372 aaatcaagagcagagattgaacaacgaagcactgattgataaaccctgctacagaagca 4431
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 541 AAGTCAAGAGCAGAGATAGAACAGAAACGACTGATGGAAGAACTGGCTACTGAAGCA 600
QY 4432 aaactaatcatcttagatataatagatgtgtgttcagaccgtttctgtaacyaatcc 4491
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 601 AACCTCATCATCTTAGACACGCTAGAGATCATCGTCCAGACTGTGTCTGTAAACGAATCC 660
QY 4492 aaagtagagcatctctgtgtgagtgctaaaagtctactactacacagcatggtcgttaacc 4551
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 661 AAAGAGAGTATCCTGGGTGTGTGCTAAAGTGTCTTACAGAGCATGGCTTGCAACCAA 720
QY 4552 agtgcagtttatctacaacactgttttgcacacagagagccttggttcaaaagtttcc 4611

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 721 AGTGAGTCTATCTGCAGCACTGCTTGCCACGAGAGAGCCCTGCTCTCAAGTTTCT 780
QY 4612 gaactctatttgaagaagacagacagcagtggtgtgattatgctcaggtctccga 4671
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 781 GAGCTCTTGTGGAAGAGAGACAGAGCAGTGTGCGATCTGTGCTCCGGCTTCCGT 840
QY 4672 cactgtagcagtagcatcgtgtacatacaggttcaacccccagtgctcccttactacta 4731
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 841 CATTGCAGTAGCAGCATCAGTACATCCGCTCGCACGCTAGTGCCTCCCTTACCTCCTC 900
QY 4732 atgagcaaaactttgagattggaataactttgccagggttaaatgcaggtaccatg 4791
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 901 ATGAGGAGCACTTTGAGATCGGGAATTAATTGCCAGAGTGAATAATCCAAGTGACGATG 960
QY 4792 tcaatctcctctgtgtggtgacacatctcagaattttaatgaagaattcttaagacgttct 4851
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 961 TCACTGTCTCTTTGGTGGGTAGCTGTCAAGATTTTAATGAAGAAATTTCTTAAGACGCTCT 1020
QY 4852 ctaagaactatattgacatatgtctgaagaagatctggaattgagggaacaacatttct 4911
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1021 CTAAGAACTATTTTGACATATGCTGAAGAGATCTAGAAATGAGGGAACACATTTCT 1080
QY 4912 gatcaggtccagatctggtttcaatctccatatgattcttctgatactgtgaaatg 4971
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1081 GATCAGGTCCAAGATTGCTCTTCAATCCATATGATGATCCCTTCTGACACTGTTAAGATG 1140
QY 4972 aaggaacaccagagagatccctgaatgtgtatgtatctaatgtacagaatttgcgaaggt 5031
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1141 AAGGAGCATCAGGAAGACCCAGAAATGCTGATGTATCTCATGTACAGAAATCGCCAAAGGGC 1200
QY 5032 taccagacctctccaga--gcatgtgacctgtgttcagaaacatgycaggaagcactca 5088
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1201 TACCAGACGCTCTCCAGACCTGCGCCTGACCTGTTGCAGAACATGGCTGGCAAGCACTCT 1260
QY 5089 gaacgaagcaatcatctgtgaagctgcacagtgcttagtccactcagcagc 5138
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1261 GAACGAAGCAATCATGCCGAAGCTGCTCAGTGCTGCTGCTCCTGCTCCTCCGACGC 1310

RESULT 3
BF309712 898 bp mRNA linear EST 21-NOV-2000
LOCUS 601891854F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:4137732 5',
DEFINITION mRNA sequence.
ACCESSION BF309712
VERSION BF309712.1 GI:11257158
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 898)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC

cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LILCM1045 row: k column: 13
High quality sequence stop: 732.

location/Qualifiers
1.898
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4137732"
/clone_lib="NIH_MGC_17"
/tissue_type="rhabdomyosarcoma"

```

/lab_host="DH10B (phage-resistant)"
/note="Organ: muscle; Vector: pOTB7; Site_1: EcoRI;
Site_2: XhoI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Size-selected
for average insert size 1.8kb. Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
```

Query Match	9.98;	Score 678.8;	DB 10;	Length 898;
Best Local Similarity	92.58;	Pred. No. 6e-146;		
Matches 757; Conservative	0;	Mismatches 57;	Indels 4;	Gaps 4;

OY	4152	glttgaacgaatgaataagcttgaccctttaagaaatcaaaaagacatgtagagcacaagcttga	4211
Db	2	GTTTGAACGAATGAATAAGCTTGACCCTTTAAGAAAATCAAAAGACATGAGAGCAAAGCTTGA	61
OY	4212	agaagctattcttgggagcatagtgtccaggccaagaaatgttacgycgaaagccgagaca	4271
Db	62	AGAAGCTATTCTTGGGAGCATAGGTGCCAGGCAGCAAGAAATGGTAGCGGGAAGCCGAGGACA	121
OY	4272	gctcgagagaagcccactctggaaglygcctttgggaagtcaaaaaatltagrtygaggaa	4331
Db	122	GCTCGAGAGAGAAGCCCACACTCTGGAAGTGCCCTTTGGAAGTCAGAANAATTTGAGGTGGAGGAA	181
OY	4332	agatatgactcactgycgtcaaaaacacacagagaagcttgcacaatlcaagagccaagattga	4391
Db	182	AGATATGACTCACTGGCGCTCAAAACACAGAGAAGCTTGACAATCAAGAGCAGAGATTGA	241
OY	4392	acacgaaagcactgatytatygaaaacctggtcacagaagcaaacctaatcatltaagatac	4451
Db	242	ACACGAAGCAGCTGATGTGATGGAAACCTGGCTACAGAGCAAACTTAATCATTTTGATGATAC	301
OY	4452	attagagattgttgttcagaaccglttctgtlaacggaatccaaagagacatctgtgtg	4511
Db	302	ATTAGAGATTG-TGTTGAGACCGTTTCTGTAAACGGAATCCAAGAGAGCATTTCTGGTGG	360
OY	4512	agtgcataaagtgtctaactacacagcatgugcctgtaaccaaaglygcagtttatctacaaca	4571
Db	361	AGTGCTAAAGTGCTACTACACAGCATGGCCGTAAACCAAGTGCGAGTTATCTACACA	420
OY	4572	ctgttttgcacacagagagccttgtttccaagtltcctgaactcttatlttgnaaaga	4631
Db	421	CTGTTTTGCTACACAGAGACCTTGTTTCAAAAGTTTCTGAACTCTTATTTGAGAGAGA	480
OY	4632	gacagaagcagtggtgctgattatgacctcagcgtctccgcacactytagcagtagcatcg	4691
Db	481	GACAGAGCAGTGTGCTGATTTATG CCTCAGGCTTCTCCGACACTGTAGCAGTAGCATCGG	540
OY	4692	tacaatacgtcacacccccagtgccctcccttaccctaataatgagcacaactttagat	4751
Db	541	TACAATAACGGTACACAGCCGCGCTCCCTTACCCTAATGAGGCAAAAAC- TTGAGAT	599
OY	4752	tgggaataacttggccagggttaaaaatgcaggtaccaatgtcaactatccctcctgtgtgg	4811
Db	600	CGGGAATAACTTTGCCAGGGTTAAAAATGCAGGTAAACAATGTCACTATCCCTCC-TGGTGGG	658
OY	4812	cacatctcagaattttaaibgaagaatlcttaagacgttctctaaagactatatltgaca-t	4870
Db	659	CACATCTCAGGAATTTTAATGAAGAATCTTAAGACGCTTCTCTAAGACTATTATTGACACT	718
OY	4871	atgctgaagaagatctggaattgaagggaacaacaacatltcctgtatcaggtccaggaatcyg	4930
Db	719	ATGCTGGAGAAGATCTGGCAATGGCGGGGAGGACACCTTGCCGGGATCAAGGCCAGAGATCGG	778
OY	4931	tttccaatctccatatgatcttctctgatactgtgaaa	4968
Db	779	GTTTCACACCACTATGAGTTTCTGGTCACTGTGAGAGA	816

RESULT	4				
BC018075					
LOCUS	BC018075	1873 bp	mRNA	linear	HTC 06-DEC-2001
DEFINITION	Homo sapiens, clone IMAGE:4792506,		mRNA.		
ACCESSION	BC018075				
VERSION	BC018075.1	GI:17390162			
KEYWORDS	HTC.				
SOURCE	human.				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1 (bases 1 to 1873)				
AUTHORS	Strausberg,R.				
TITLE	Direct Submission				
JOURNAL	Submitted (03-DEC-2001) National Institutes of Health, Mammalian				

REMARK	NIH-MGC Project URL: http://mgc.ncl.nih.gov
COMMENT	Contact: MGC help desk

Email: cgapbs-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shitaki
Toshiyuki and Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www.shgc.stanford.edu>
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
Series: IRAK Plate: 32 Row: b Column: 21

This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein

This clone has the following problem: frame shifted.
Location/Qualifiers

```
location/Qualifiers
1. .1873
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4792506"
/tissue_type="Brain, hypothalamus"
/clone_lib="NIB_MGC_96"
/lab_host="DH10B"
/note="Vector: pBluescript"
```

```

/note="Vector: pBluescript"
BASE COUNT      604 a      317 c      365 g      587 t
ORIGIN

```

Query Match	9.98;	Score 678;	DB 11;	Length 1873;
Best Local Similarity	100.08;	Pred. No. 1.2e-145;		
Matches 678;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	5498	agcggatgtl	ttggcacctat	tttcgtgtgt	gttctta	ggaaccaag	ttcgggatt	ttgg	5557
Db	1178	AGCGGATGTT	TGGCACCTAT	TTTCGTGTGT	TTTATGGA	ACCAAGTTC	GGGGATTGG		1237
QY	5558	atgaacaaga	atttgtttaca	agagcctgc	ataaccaact	tcgagagata	tatctaca		5617
Db	1238	ATGACACA	GATTTGTTT	TACAGGAG	CGCTGCAAT	AACCAACTT	GCAGAGAT	ATCTCACA	1297
QY	5618	gattgagagg	atttaccgga	gaagaatt	ttgagagat	gtgtgaag	taatacaag	act	5677
Db	1298	GATTGGAG	GGGATTTT	TACGGAGAA	GATTTTG	GAGAGGAT	GTGGTTGA	AGTAATCAAGACT	1357
QY	5678	ctaatcctg	tagacaagt	gtcaaat	tagatccta	acaagycata	tattcagatt	accatg	5737
Db	1358	CTAATCCT	GTAGACA	AGTGTGTA	ATTAGATC	CTAACAGC	ATATATT	CAGATTACTATG	1417

QY	5738	tgagagccatactttgacacatatgagatgaagacagaaatcacctattcgacaaaaatt	5797
Db	1418	TTGAGGCCATTACTTTGACACATATGAGATGAAGACAGAACATCACTTATTCGACAAAAATT	1477
QY	5798	acaatctcgtcgatcatgttactgtacaccccttactttaagatggccgtgcccattggg	5857
Db	1478	ACAATCTTCGTCGATTCATGTACTGTACACCCCTTACTTTAGATGGCCGTGCCCATGGGG	1537
QY	5858	aacttcataaacaatccaagaagaagaccattctgactaacgtctcatatgcctttccttata	5917
Db	1538	AACATCATGAACATATCAAAAGAGAACCATTTCTGACTACGCTCATGCGCTTTCCTTATA	1597
QY	5918	ttaaaacaagggtcaatgttcactcataaagaagagatcatcttaacaccaattgaagttg	5977
Db	1598	TTAAACAAGGGTCATATGTCACATATAAGAAGAGATCATCTTAACACCAATTGAAGTTG	1657
QY	5978	ctattgaaggacatgcagaaaaaacacacaggaagttggcatttgcacaacatcagatcccg	6037
Db	1658	CTATTGAGGACATGCAGAAAAAGACACAGAGATTGGCAATTTGCCAACACATCAGGATCCCG	1717
QY	6038	cagaccccaaaaatgtcttcagatggtactccaggatctgtagccaccacagtgaaatcagg	6097
Db	1718	CAGACCCCAAAATGCTTCAGATGGTACTCCAGGATCTGTAGGCACACAGATGAATCAGG	1777
QY	6098	ggccttgggaagttgccaggtttttctgtctgaaatacctaagtgaccccaagctcttca	6157
Db	1778	GGCCTTGGGAAGTTGCCCAGGTTTCTGTCTGAATACCTAGTGACCCCAAGCTCTTCA	1837
QY	6158	gacatcataataaactgc	6175
Db	1838	GACATCATATAATAACTGC	1855
RESULT	5		
LOCUS	BC018076	1873 bp	mRNA
DEFINITION	Homo sapiens, clone IMAGE:4792507,		linear
ACCESSION	BC018076		HTC 06-DEC-2001
VERSION	BC018076.1	GI:17390164	
KEYWORDS	HTC.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 1873)		
TITLE	Strausberg,R.		
JOURNAL	Direct Submission		
	Submitted (03-DEC-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA		
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov		
COMMENT	Contact: MGC help desk		
	Email: cgabs-r@mail.nih.gov		
	Tissue Procurement: Miklos Palkovits, M.D., Ph.D.		
	CDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki Toshlyuki and Piero Carninci (RIKEN)		
	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)		
	DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305		
	Web site: http://www-shgc.stanford.edu		
	Contact: (Dickson, Mark) mcdépxil.stanford.edu		
	Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.		

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov

Series: IRAC Plate: 32 Row: B Column: 20

This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein

This clone has the following problem: frame shifted.

FEATURES

Location/Qualifiers

source	1. 1873
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/clone="IMAGE:4792507"
	/tissue_type="Brain, hypothalamus"
	/clone_lib="NIH_MGC_96"
	/lab_host="DH10B"
	/note="Vector: pBluescript"
BASE COUNT	604 a 317 c 365 g 587 t
ORIGIN	

Query Match 9.9%; Score 678; DB 11; Length 1873;
Best Local Similarity 100.0%; Pred. No. 1.2e-145;
Matches 678; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	5498	agcggatgtttggcacctatttctgtgtgttttatgtgaaccaagttcgggattcgg	5557
Db	1178	AGCGGATGTTGGCACCTATTTCGTGTGGTTTATGGAACCAAGTTCGGGATTTGG	1237
QY	5558	atgaacaagaattgtttacaagagcgctgcataaaccacaaattgcagagatatctcaca	5617
Db	1238	ATGAACAAGATTGTTTACAGAGAGCGCTGCATAAACCAACTTGCAGAGATATCTCACA	1297
QY	5618	gattggaggatllttacggaagaagatttggagagatgtgtgtgaagtaatcaagaact	5677
Db	1298	GATTGGAGGATTTTACGGAGAAGATTTGAGAGAGATGTGTTGAAGTAATCAAGACT	1357
QY	5678	ctaatcctgtagacaagtgtaaatagatcctaacaagcatabatlacagattaccatg	5737
Db	1358	CTAATCCTGTAGACAAGTGTAAATTAGATCCTAACCAAGGCATATATTAGATTACCTATG	1417
QY	5738	tgagagccatactttgacacatatgagatgaagacagaatcacctattcgacaaaaatt	5797
Db	1418	TGAGGCCATTACTTTGACACATATGAGATGAAGACAGAACATCACTTATTCGACAAAAATT	1477
QY	5798	acaatctcgtcgatcatgttactgttacaccccttactttaagatggccgtgcccattggg	5857
Db	1478	ACAATCTTCGTCGATTCATGTACTGTACACCCCTTACTTTAGATGGCCGTGCCCATGGGG	1537
QY	5858	aacttcataaacaatccaagaagaagaccattctgactaacgtctcatatgcctttccttata	5917
Db	1538	AACATCATGAACATATCAAAAGAGAACCATTTCTGACTACGCTCATGCGCTTTCCTTATA	1597
QY	5918	ttaaaacaagggtcaatgttcactcataaagaagagatcatcttaacaccaattgaagttg	5977
Db	1598	TTAAACAAGGGTCATATGTCACATATAAGAAGAGATCATCTTAACACCAATTGAAGTTG	1657
QY	5978	ctattgaaggacatgcagaaaaaacacacaggaagttggcatttgcacaacatcagatcccg	6037
Db	1658	CTATTGAGGACATGCAGAAAAAGACACAGAGATTGGCAATTTGCCAACACATCAGGATCCCG	1717
QY	6038	cagaccccaaaaatgtcttcagatggtactccaggatctgtagccaccacagtgaaatcagg	6097
Db	1718	CAGACCCCAAAATGCTTCAGATGGTACTCCAGGATCTGTAGGCACACAGATGAATCAGG	1777
QY	6098	ggccttgggaagttgccaggtttttctgtctgaaatacctaagtgaccccaagctcttca	6157
Db	1778	GGCCTTGGGAAGTTGCCCAGGTTTCTGTCTGAATACCTAGTGACCCCAAGCTCTTCA	1837
QY	6158	gacatcataataaactgc	6175
Db	1838	GACATCATATAATAACTGC	1855
RESULT	6		
LOCUS	AK004816	2694 bp	mRNA
DEFINITION	Mus musculus adult male lung cDNA, RIKEN full-length enriched library, clone:1200017A24:homolog to BA165F24.1.1 (NOVEL PROTEIN SIMILAR TO DROSOPHILA CG6630 AND CG11376, KIAA1058, RAT TRG (ISOFORM 1)) (FRAGMENT), full insert sequence.		
ACCESSION	AK004816		

VERSION AK004816.1 GI:12836282
KEYWORDS HTC, CAP trapper.
SOURCE Mus musculus (strain:C57BL/6J) adult male lung cDNA to mRNA,
clone_lib:RIKEN full-length enriched mouse cDNA library
clone:1200017A24.

ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS Carninci,P. and Hayashizaki,Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636

REFERENCE
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159

REFERENCE
AUTHORS 3 (sites)
Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
Konno,H., Akiyama,J., Nishi,K., Kitsuunai,T., Tashiro,H., Itoh,M.,
Suni,N., Ishi,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861

REFERENCE
AUTHORS 4 (sites)
The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
REFERENCE 5 (bases 1 to 2694)

REFERENCE
AUTHORS Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A.,
Arakawa,T., Baldarelli,R., Bono,H., Brownstein,M., Butt,C.,
Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hanagaki,T.,
Hara,A., Hayatsu,N., Hill,D., Hiramoto,K., Hiraoka,T., Hori,F.,
Hume,D., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kasukawa,T.,
Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S.,
Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K.,
Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Quackenbush,J.,
Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D.,
Schriml,L., Shibata,K., Shibata,Y., Shinagawa,A., Shizaki,T.,
Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,I.,
Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yamanaka,I.,
Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and
Hayashizaki,Y.
TITLE Direct Submission
JOURNAL Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)

COMMENT Please visit our web site (http://genome.gsc.riken.go.jp/) for
further details.
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues. First strand cDNA was primed with a primer
[5' GAGGAGAGAGCGGCGCGACTGAGTGTCTTTTCTTTTCTTNN 3'], cDNA was
prepared by using trehalose thermo-activated reverse transcriptase
and subsequently enriched for full-length by cap-trapper. Second

FEATURES
source strand cDNA was prepared with the primer adapter of sequence[5'
GAGGAGAGAGAGCTCCAGAGCTCAATTAATTAATTAACCCCCCCC 3']. cDNA was
cleaved with XhoI and SstI. Cloning sites, 5' end: SstI; 3' end:
XhoI. Host: SOLR.
Location/Qualifiers
1..2694
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="MGI:1907255"
/db_xref="taxon:10090"
/clone="1200017A24"
/sex="male"
/tissue_type="lung"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
22..1563
/note="data source:SPTR, source key:Q9H1Q3, evidence:ISS
homolog to BA165F24.1.1 (NOVEL PROTEIN SIMILAR TO
DROSOPHILA CG6630 AND CG11376, KIAA1058, RAT TRG (ISOFORM
1)) (FRAGMENT)
putative"
/codon_start=1
/protein_id="BAB23587.1"
/db_xref="GI:12836283"
/translation="MQTFPPMQVEELCNINSILYDTVKMREFQDEPMLDMYRI
AKSYQASPDRLTLWLOMAEKKTKCFTEAMCLVHAALVAEYLSMLEDSYLPVG
SVSFQNISSNVLEESAVSDTLSPDEDEGVCGRYFTESGLVGLLEAELESTGLYE
TVNEVYKLVIPILEARDFRKLSTHDKLQKAFDNIINKDHKRMFGYFRVGYGSRF
GLDDEQEEVYKEPATYKLEIPLEISHRLLEGFYGCQGAEEVEVIKIDSTPDYKPKAY
IQITFEVPEYFDEYEMKDRVTYFEKFNLRFRMTTPTPLEGRPGLHQRNRTVL
TMAEPYIKTRIRVSQKEEFVLPPIEVAIEDMKKTKQLAVATHQEPDAKMLQVLO
GSVGATVNOGPLEVAQVFLAEIPADPKLYRHNNKLRLCFKEFIMRCGEAVEKNRLIT
AEQREYQQLKKNNKLRDLSLRPMIERKIPELYKPIFRVDSQKRDSEFRSSFRKCETO
LSQGS"
polyA_signal 2677..2682
/note="putative"
polyA_site 2694
/note="putative"
BASE COUNT 752 a 646 c 682 g 614 t
ORIGIN
Query Match 9.7%; Score 662; DB 11; Length 2694;
Best Local Similarity 66.9%; Pred. No. 6.5e-142;
Matches 1006; Conservative 0; Mismatches 485; Indels 13; Gaps 4;
QY 4870 tatgctgaagaagatctggaattgaggygaacaacatttcctgatacaggtccagatctg 4929
1 TTTCAGAGAAGAGACACGGCCATGCAGACAACCTCTTTCCCATGCAGGTGAGGAACCTT 60
Db 1 TTTCAGAGAAGAGACACGGCCATGCAGACAACCTCTTTCCCATGCAGGTGAGGAACCTT 60
QY 4930 gtttcacatcctcatatgattccttcgtatactgtgaaatgaaaggaacaccagagagat 4989
1 11111111 1 11111 1 111111111111111 11111 11111 11
Db 61 CTCTGCAATCTGACACAGCATTTCTGTACGACACAGTGAAGATGAGGGAATTCCAGGAAGAC 120
QY 4990 cctgaatgttgattgatctaatgtacagaattgccaaaggyttaccagacctctccaga- 5048
11111111 1 1 111111111111111111111 1 11111 1 11 11
Db 121 CCTGAGATGCTTATGGACCTCATGTACAGAAATTGCCAAGAGGCTACCGACGCTGAC 180
QY 5049 --gcgattgacctggttcagaaacatgcaaggcaagcaactcagaacgaagcaatcatgct 5106
111 1111 111 111111111111111111 11111 1 1
Db 181 CTGCGGCTGACTTGCTCCAGAACATGGCAGAGAAACACACACTAAGAGAAGTGCTTCACA 240
QY 5107 gaagctgcacagtgctctagtcactcagcagcaactgttgctgaatataggatgctg 5166
11111 111 111 111 1111 1111 111 111 111 1111111111
Db 241 GAGGCCGCATGTGCTGTGGTCATGCAGCCGCCCTGCTGGCCGAGTACCTGAGCATGCTG 300
QY 5167 gaggaccggaatatcttcctgtgagatgtgtaacatttcagaatatcattcatatgct 5226
1111111 1 11 11 111111 1 11 1 11111111111111 11 11111
Db 301 GAGGACACACAGCTACCTGCGCGGTGGGACGCTCAGCTTTCAGAAATTTCTTCCAATGTG 360
QY 5227 ttagaagaatctgcygctcagatgatgtgtatctccagatgaagaaggtatctgctc 5286
1 11 11 11111 11111 11111 1 11 11 11111 11 11 111111
Db 361 CTTGAGGAGTCTGCACTCTCTGATGACACCTTTGTCACCTGATGAGGACGGCGTATGCTCT 420

QY	5287	ggaaaatacttctactgagtcaggacttgtggtatctactgtaacaagaocgtgtctctc	5346
Db	421	GGTCGGTACTTCACTGAGAGTGGCCCTGGTGGCCCTCCTGAGCAGAGCTGGAGCTCTTC	480
QY	5347	tctatgctgcatgtatgaagcagttatgaagtttacaagttacttatctctatcat	5406
Db	481	AGCACGGGAGGCTGTACGAGACGGTTAATGAAGTCTACAAAGCTGTCATCCCTATCTTG	540
QY	5407	gaagctaactcgygatgtaagaanaactatccacaattcactgtaactcaagaagcatc	5466
Db	541	GAGGCACACAGAGATTTCCGGAAGCTGACCTCCACTCAGACACAAGCTGCAGAAGGCCCTTC	600
QY	5467	agcaaaatgttcatacagagtaactgctgctggagcggatgttggcacctatttctgtgt	5526
Db	601	GATAACATCATCAACAAG-----GACCATAAGAGGATGTTTGGGACCTACTTCCGAGTT	654
QY	5527	ggttttatgtaaccaagtctcggggatttggatgtaacaagaatttgtttacaaggacct	5586
Db	655	GGTTTCTACGGATCCCGATTTGGGGATTGGATGAGCAGGAGTTCGTGTACAAAGAACCC	714
QY	5587	gcaataaccaactctgcagagatatctcacagatgtaagggaatttcacgagaagaattt	5646
Db	715	GCAATCACGAAGCTCCCGGAGATCTCACATAGACTAGAGGGATTTTATGGCCAGTGTTC	774
QY	5647	ggagaggaatgtgttgtaagtaatcaaaagctctaactcgttagacaagtgtaaattagat	5706
Db	775	GGTCAGAGTGTGTGAAGTGATAAAGACTCTACTCCAGTGAGCAAAAACCAAGTTGAT	834
QY	5707	cctaacaaggcatatattcagatllacatattgtggagcatacttggacatatgagatg	5766
Db	835	CCTAACAGGCCCTACATTCAGATCACTTTGTGAGCCTTACTTTGATGAATATGACATG	894
QY	5767	aagcacagaatcacctatttcgacaaaattacaatcttcgtcgaatctcatgtactgtaca	5826
Db	895	AAAGACCGGGTGACCTACTTCGAGAAGATTTCAACTCCGGAGGTTTCATGTACACCACC	954
QY	5827	cccttactttagatggccgtgcccattggygaacttcatgaacaattccaagaagacc	5886
Db	955	CCCTTCACCCCTGAGGGGAGACCCCGGGCGAGCTTCATGAGCAACACCCGAGAACACC	1014
QY	5887	attctgactacgtctcatcactgcttctcctatatattaaacaaggtcgaatgtcactataa	5946
Db	1015	GTGCTCACCAACCATGCGAGCCCTTCCCTACATCAAGACCAGGATCCGAGTCAGCCAGAAA	1074
QY	5947	gaagagatcatcttaacaccaaattgaaatgtctatgtaggacatgacagaaaagacacag	6006
Db	1075	GAGGAGTTCGTTTGACTTCGATTGAGGTGCGCATTTGAAGATATGAAGAAGAACCCCTG	1134
QY	6007	gaattgycatttgcacaacatcaggatcccgacagaccccaaaatgctcagatggtactc	6066
Db	1135	CAGTTAGCCGTGGCCACTCACCAAGAGCCCTGATGCCAAAGATGCTGCAAAATGCTACTG	1194
QY	6067	cagggaatcgttaggcacacacagtgaaatcaggggccttggaaatgtgcccagttttctg	6126
Db	1195	CAGGGCTCTGTAGAGCCACTGTAATTCAGGGACCACTGGAGGTGGCCCAAGTGTCTTG	1254
QY	6127	tctgaatatcctagtgacccaagaagctcttcagacatcataataactgycactgcttt	6186
Db	1255	GCTGAATTCACAGCTGACCCAAAGCTTACCAGCATCAACAACAAGCTGAGGTGTGCTTC	1314
QY	6187	aaagatttactaaaaaggtgtgaagatgcttcctaagaaaaaataagagcttaattggyccg	6246
Db	1315	AAGGAGTTCAATAATGCCATGCGGAGAGAGCCCGTGGAGAAGAAGAGGCGACTCATCACCGCA	1374
QY	6247	gttcaaaaagagtatcaaaaggaattggyga--aactatcttcgccttaagaagggccta	6304
Db	1375	GAGCAGCGGGAGTACCAAGCAGGAGCTGAAGAAGAACTACAACAAGCTGAGAGACAGCCTC	1434
QY	6305	cagccctagat--cacagaaagtcacctcagttatccaagccagtatgtgttgcctcgc	6362
Db	1435	AGGCCATGATGTAGCGGGAATAATCCACAGAGCTTACAAAGCCATATTCAGAGTTGACAGT	1494

QY	6363	caca	6366
Db	1495	CAGA	1498

RESULT	7				
LOCUS	BI088758	795 bp	mrna	linear	EST 20-JUN-2001
DEFINITION	602851547f1 NIH_MGC_10 Homo sapiens CDNA clone IMAGE:4993330 5', mrna sequence.				
ACCESSION	BI088758				
VERSION	BI088758.1	GI:14507088			
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				

REFERENCE AUTHORS TITLE JOURNAL COMMENT	
1 (bases 1 to 795) NIH-MGC http://mgc.ncbi.nlm.nih.gov/ . National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Robert Strausberg, Ph.D.	

COMMENT

Contact: Robert Strausberg, Ph.D.
Email: cgapsb-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: Incyte Genomics, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: L1AM11013 row: m column: 11
High quality sequence stop: 640.

FEATURES	Location/Qualifiers
source	1..795

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4993330"
/clone_lib="NIH_MGC_10"
/cell_line="MGC36"
/lab_host="DH10B"
/note="Organ: cervix; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.5 kb. Library prepared by Life
Technologies."

```

Query Match	9.5%;	Score 651.4;	DB 10;	Length 795;
Best Local Similarity	96.0%;	Pred. No. 1.3e-139;		
Matches 765; Conservative	0;	Mismatches 21;	Indels 11;	Gaps 9;

QY 4074 ctccagtccttcgacgtaaacgcgc-tattagatctgctcttaactctgtgtctcttcttg 4132

QY 4133 agtataaaggggaaaaaagtgttttgaaagaaatgaatagcttgacctttaagaatccaag 4192
|||||
Db 61 AGTATAAAGGCAAAAAGTGTTTGAACGAATGATAGCTTGACCTTTAAGAATCAAAAG 120

Qy	4193	acatgagagcaaaagc	ttgaagaagctatct	cttggagcatagtg	ccagcgcaagaatgg	4252
Db	121	ACATGAGAGCAAAAGCTT	GAGAAGCTATTCTT	GGAGCATAGCTGCC	AGGCAAGAAATGG	180

Qy 4253 tacggcgaagccgagacagctcgagagaagcccatctgtaagtgcctttggaagtcgaag 4312
|||||
Db 181 TACGGCGAAGCCGAGGACAGCTCGAGAGAAGCCCATCTGGAAGTGCCCTTGGAAAGTCGAAG 240

Oy 4313 aaatttgaagtgtagaaagatatgactcactgcyctcaaacacagagaagtgcaca 4372
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 241 AAAATTGAGCTGGAGGAAGATATGACTCACTGGCGGTCAAACACACAGAGAAGCTTGACA 300

QY 4373 atcaagagcagagattgaaacacgaagcactgtattgatgaaacctggtacagaagcaa 4432

|||||
Db 301 AATCAAGACGACGAGATTGAACACGACGACTGATGTGGAACCTGGCTACAGAACCA 360
QY 4433 accatacatttagatacatagagattgtgttcagaccglttctgttaacggaatcca 4492
|||||
Db 361 ACCTAATCATTTTAGATACATTAGAGATTG-TGTTACAGCCGTTTCTGTAAACGGAATCCA 419
QY 4493 aagagagcattctgtgtgagtgctaaaagtgtctactacacagcatggcctgttaaccaa 4552
|||||
Db 420 AAGAGACATTCTTGGTGAGTCTAAAGTGCTACTACACAGCATGGCCTGTAAACCAA 479
QY 4553 gtgcagtttatctacaacactgtttgtctacacagagcctgtgttcaagtctcg 4612
|||||
Db 480 GTGCAGTTTATCTACACACTG-TTTGCTACACAGAGAGCCTTGTTCAAGTTCCTG 538
QY 4613 -aactctatttgaagaagagacagcagctgtgtctgatttatgcctcagcgttcocga 4671
|||||
Db 539 AAACCTCTTATTGAAGAAGACAGACAGCAGTGTGCTGATTTATGCCTCAGGCTTCGCCA 598
QY 4672 cactgtacga-gtagcatcggtacaatacagtgtaacacccccagtg-cctcccttacctac 4729
|||||
Db 599 CACTGTAGCAGCTAGCATCGGTACATAACGGTACACACGCCAGTGACGTCCTTACTAC 658
QY 4730 taatgagcaaaaactttagattggaat-aacttgcaggggttaaatgca--ggtac 4786
|||||
Db 659 TAATGAGCAAAACCTTGAGATGGGGAATAAAGTTGCGAGGTTCAACATGAACGGTAA 718
QY 4787 caatgtcattatcctcctgtgtggtgcacatctca--gaatttaatgaagaattctaaag 4844
|||||
Db 719 CAATGTCACTATCCTCCTTGTTGGTGCGACATCTCAAGAGATTTAATGAAGAATTTCTACG 778
QY 4845 acgttctctaaagacta 4861
|||||
Db 779 AAGTACTCTAAGCAAA 795

RESULT 8
AK014226 1051 bp mRNA linear HTC 19-JAN-2002
LOCUS
DEFINITION Mus musculus 13 days embryo head cDNA, RIKEN full-length enriched library, clone:3110056M06.homolog to BA165F24.1.1 (NOVEL PROTEIN SIMILAR TO DROSOPHILA CG6630 AND CG11376, KIAA1058, RAT TRG (ISOFORM 1)) (FRAGMENT), full insert sequence.
ACCESSION AK014226
VERSION AK014226
KEYWORDS AK014226.1 GI:12851941
SOURCE HTC; CAP trapper.
Mus musculus (strain:C57BL/6J) 13 days embryo head cDNA to mRNA, clone_lib:RIKEN full-length enriched mouse cDNA library
clone:3110056M06.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (sites)
AUTHORS Carninci,P. and Hayashizaki,Y.
JOURNAL High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE 2 (sites)
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
REFERENCE 3 (sites)
AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwaki,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,

TITLE Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
JOURNAL RIKEN integrated sequence analysis (RISA) system--384-format
MEDLINE sequencing pipeline with 384 multicapillary sequencer
PUBMED Genome Res. 10 (11), 1757-1771 (2000)
REFERENCE 20530913
11076861
4 (sites)
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
AUTHORS Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A., Arakawa,T., Baldarelli,R., Bono,H., Brownstein,M., Bult,C., Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hill,D., Hiramoto,K., Hiraoka,T., Horii,F., Hume,D., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Quackenbush,J., Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Schriml,L., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yamanaka,I., Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
TITLE Direct Submission
JOURNAL Submitted (10-JUL-2000) Yoshinide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Gemomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
COMMENT please visit our web site (http://genome.gsc.riken.go.jp/) for further details.
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Gemomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGAGAGAGAGATCCAGAGAGCTCTTTTTTTTTT 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 50.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCGAGTTAATTAAATTAATCCCCCCCCCCC 3']. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI, 3' end: SstI. Host: SOLR.
FEATURES
source Location/Qualifiers
1. 1051
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="MGD:MGI:1896359"
/db_xref="taxon:10090"
/clone="3110056M06"
/tissue-type="head"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="13 days embryo"
3. 653
/note="data source:SPTR, source key:Q9H1Q3, evidence:ISS homolog to BA165F24.1.1 (NOVEL PROTEIN SIMILAR TO DROSOPHILA CG6630 AND CG11376, KIAA1058, RAT TRG (ISOFORM 1)) (FRAGMENT)"
putative"
/codon_start=1
/protein_id="BAB29215.1"
/db_xref="GI:12851942"
/translation="MKDRITYFDKNYNLRRMYCTPFTLDGRAHGLHEQFKRKTILT TSHAFPIKTRNVNTHKEIILTPIEVAIEDMQKKTQELAFATHQDPADPKMLQVLIQ GSVGTVNQGPLEVAQVFLSEIPGDKLFRHHNKKRLRCFDPFTKCEDALRKNSLIQ PDOKEYQRELERNYHRLKEALOPLINRKIPQLYKAVLPVCHRDSFSRMSLRKML"

polyA_signal 1031. .1036
/note="putative"
polyA_site 1051
/note="putative"
BASE COUNT 337 a 207 c 199 g 308 t
ORIGIN

Query Match 9.5%; Score 646; DB 11; Length 1051;
Best Local Similarity 82.6%; Pred. No. 2.4e-138;
Matches 876; Conservative 0; Mismatches 160; Indels 24; Gaps 11;

QY 5763 gatgaagcacagaatcacctatttcgcagcaaaaattacaatcttcgtcattcatgtactg 5822
|||||
Db 2 GATGAAGCACAGATACACTTATTTGGACAAAAATTATATCTCCGGCGTTTCATGTACTG 61
QY 5823 tacacccttacttagatgycgcgtgcccatggggaacctcatgaacaattcaaaaggaa 5882
|||||
Db 62 CACACCCCTCACTTTAGATGGCCGCTGCTCATGGGAGCTTCATGAACAATTCAAACGGAA 121
QY 5883 gaccattctactacgtctcatgaccttccttataataaacaagggtcaatgtcacitca 5942
|||||
Db 122 GACCATTTCTGCAACAGCTCTCATGCTTTCCTTACATTAAACAAGATCAATGTCACTCA 181
QY 5943 taagaagagatcatcttaacaccaattgaagtgtatttgagacatgcagaaagac 6002
|||||
Db 182 CAAGAAGAGATTAATCTTAACACCAATTGAAGTTGCTATCGAAGACATGCAGAAAAAGAC 241
QY 6003 acaggaatgycatttgcacacacatcagatcccgacagccccaanaatgcttcagatggt 6062
|||||
Db 242 TCAGGAATGGCATTTGCAACACATCAGGATCCAGACAGCCCAAAATGCTCAAAATGCT 301
QY 6063 actccaggatctgtagcaccacacagtgaatcagggcccttgggaagttgccaggtttt 6122
|||||
Db 302 CCTCCAGGGGTCTGTAGGAACAACACTGTAACCAAGGAGCCTCTGAGGTTGCCCGACTT 361
QY 6123 tctgtctgaatatacctagtgcaccaaaagctcttcagacatcataataaactgcgactctg 6182
|||||
Db 362 CCTGTGAAATACCTGGTGACCCGAAGCTCTTCAGACATCACAACAAGCTGCCCTTTG 421
QY 6183 ctttaagaatttactaaaagtggtgaagatgaccttaagaaaaataagagcttaattg 6242
|||||
Db 422 TTTCAAGCACTTACTAAAGGTGTGAGGATGCCCTTACGAAAAAATAAGAGCTTAAATTGG 481
QY 6243 gccggttcaaaaaggagatcaaaagggaattg--gggaaactatctcgacctaaagagc 6300
|||||
Db 482 ACCAGATCAAAAAGAGATCAAGAGAAGAACTGGAGAGAAATTACCATGCCCTTAAAGAGGC 541
QY 6301 cctacagccctagatca--cagaagaagtcctcagttatccaagccagtatgtgtccc 6358
|||||
Db 542 TCTGCAGCCCACTGATAAACAGAAAAATCCCTCAGCTATACAAAGGCTGTGTGCTGTAC 601
QY 6359 ctgccacagagatcctcctcagtcgaatgagcttgcgcaaaatgatactctaactgaatg 6418
|||||
Db 602 TTGCCACAGAGATTCTTACGCCGAATGAGCCTTCGAAAAATGGAACCTCTAAAT--ACA 659
QY 6419 cacttgtttatcatctctgcagaagagccatgtatccaacatcgagtgtgaaaagatctat 6478
|||||
Db 660 CACTGTCTCTATTCATTTGAAGAAGAACCATGTATTCACACTGAGTGTG--AGGCTTAT 716
QY 6479 tggaaaccaaca-tggaaatggaattctggaattatattcattgaagaatgcagtgcc 6537
|||||
Db 717 TGAATAATAGGACTTGGATGTCAATCTGGAAGCAATAGT--ATTGATTTACTCTCTGAAG 774
QY 6538 aagaataatacaaatgtagatgtttaaagcgttgagaatcatggtcatgtgttctaagt 6597
|||||
Db 775 AATGTGCTCCAGATGCAGGCTGTT-ACGTGTGAGGCTTGTGGC-GTGGTTTCTAAACCTT 832
QY 6598 cgggtaacaagctgtatctt---ttaagacatttaatgactcaagaagtacaactaac 6653
|||||
Db 833 CTGTTAACATGTGTCTATCTTAAAAACAAGGTTTAAATGATTCAAGGTTACAGTACAC 892
QY 6654 attaccatattatatacatagctaaggttaaaaaatttaccatttaagttcgtattt 6713

Db 893 ATTTACAATTATTATACCATAGCTAATGTT-AAATTATTACTTTTAAGTCTTGT 951
QY 6714 tttaattatataccacattatagattca-----tttggaccattttaaatgtagtaa 6768
|||||
Db 952 TTTAATTATATATACATTATAGGTCATTCATTTTTGAACCATTTTAAATGATAGTAA 1011
QY 6769 tgcctattttaagyractaaaaatatgtgaatgtttac 6808
|||||
Db 1012 TGCTTATTTTAAAGTACTATTAAATATGTGAATGTTTAC 1051

RESULT 9
BG247108 931 bp mRNA linear EST 13-FEB-2001
LOCUS
DEFINITION 602359776F1 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:4488235 5',
mRNA sequence.
ACCESSION BG247108
VERSION BG247108.1 GI:12756923
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 931)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10334 row: 9 column: 20
High quality sequence stop: 728.

FEATURES
source location/Qualifiers
1..931
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4488235"
/clone_lib="NCI_CGAP_Mam1"
/tissue_type="tumor, biopsy sample"
/dev_stage="10 months, virgin"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"

BASE COUNT 275 a 198 c 228 g 230 t
ORIGIN

Query Match 8.9%; Score 610.4; DB 10; Length 931;
Best Local Similarity 90.3%; Pred. No. 4e-130;
Matches 688; Conservative 0; Mismatches 66; Indels 8; Gaps 3;

QY 5244 ctcaagatgattgtgtatctccagatgaagaaggtatctgctctggaataacttactga 5303
|||||
Db 1 CTCAGATGATGTGTATCTCCAGATGAAGAAGCATTTGCTCTGGAAGTACTTCACCTGA 60
QY 5304 gtcaggactgtggaatctactggaacaagcagctgcttcctctctataagctggcatgta 5363
|||||
Db 61 ATCGGACTGTGGCTTACTTGGAACAAGCGGCTGCTTCTCCATGGCGGCATGTA 120
QY 5364 tgaagcagtaatgaagttacaagttactatctcattcatgaagcgaatcgggatgc 5423
|||||
Db 121 CGAAGCAGTTAATGAGGTTTACAAAGTACTCATTCCTATTCATGAAGCTAATCGGATGC 180

QY	5424	aaagaaactatccacaattcaltgyltaaaccttcaaagaagcatcagcaaaattgttcatca	5483
Db	181	AAAGAACTCTCCACAATTCATGGCAAACTTCAGGAGGCATTCAGCAAAATCGTCCATCA	240
QY	5484	gaagtaactggtcggagcgcgatgtttgycacctatttcgtgtgtgttttatggaacca	5543
Db	241	G-----GATGCTAAGCGGAGTGTGGCACCTATTTCGTTGTGTTTATGGAACCAA	294
QY	5544	gttcgggatttggatgacaagaattgtttacaagagccctgcataaccaacttgc	5603
Db	295	GTTCCGGGATTTGGATGAACAGGAATTTGTTTACAAGAGAGCCCTGCAATAACCAAACTCGC	354
QY	5604	agagatattctcacagattggagggttttaacgagaaagatttggagagatgtgttga	5663
Db	355	AGAGATTTCTCACAGATTTGGAGGGAATTTATGGAAGAAGATTTGGAGAGATGTGCTTGA	414
QY	5664	agtaatcaaagactctaatacctgttagacaagtgtlaaattagatccctaacaagcatatat	5723
Db	415	AGTAATCAAGGACTCTAATCCCTGTAGACAAGTGCAAAATTAGATTCAAACAAGGCATATAT	474
QY	5724	tcagattacctatgtlgyagccatactttgacacacatatagagatgaaggacagaatcaccta	5783
Db	475	TCAGATTACTACGTGGAGCCCTTCTTTGACACCTATGAGATGAAGAGACAGAATCACTTA	534
QY	5784	tttcgacaaaattacaatcttcgtlccagatcaltgtaactgttacacccttacttagatgg	5843
Db	535	TTT-GACAAAATTAATATCTCCGCGTTTCATGTACTGCACACCCTTCACTTAGATGG	593
QY	5844	ccgtgcccatggygggaactlcatatgnaacaattcaaagaagaccattctgactacgtlcca	5903
Db	594	CCGTGCTCATGGGGAGCTTCATGAACAATTCAAACGGAAGACCATTCTGACAACGCTCA	653
QY	5904	tg-cccttcctatatataaacaagygltcaatgtctacataaagaagatcatcttaa	5962
Db	654	TGTCCTTTCCTTACATTAAACAAGAGTTCATGTCACTCACAAAGAAGAGATTATCTTAA	713
QY	5963	caccaatgaagtgctcalttgagacatgcagaaaaagacac	6004
Db	714	CACCAATTGAAGGTGCTATCCGAAGACTGCAGAAACAGACCC	755
RESULT	10		
BE891826		958 bp	EST 20-OCT-2000
LOCUS	BE891826		
DEFINITION	601434272F1 NIH_MGC_72 Homo sapiens CDNA clone IMAGE:3919720 5',		
ACCESSION	BE891826		
VERSION	BE891826.1	GI:10351537	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	1 (bases 1 to 958)		
JOURNAL	NIH-MGC http://mgc.nci.nih.gov/ .		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)		
	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgapbs-r@mail.nih.gov		
	Tissue Procurement: ATCC/DCTD/DTF		
	CDNA Library Preparation: Life Technologies, Inc.		
	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)		
	DNA Sequencing by: Incyte Genomics, Inc.		
	Clone distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LLNL at:		
	http://image.llnl.gov		
	plate: LIAM9749 row: 0 column: 17		
	High quality sequence stop: 637.		
FEATURES	Location/Qualifiers		
source	1..958		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:3919720"		

Query Match	Best Local Similarity	8.8%;	Score 599.4;	DB 10;	Length 958;
Matches 742;	Conservative	0;	Mismatches 101;	Indels 22;	Gaps 6;
Db	1	GGAAATTGAGGGAACAACATTCTCGATCAGGTCCAGGATCTGGTTTCAATCTCCATAT	60		
QY	4887	ggaattgagggaaacaacatttcctgatcaggtccagagatctgttltcaatcctcatat	4946		
Db	4947	gattcttctgatactgtgaaaaatgaagaaacaccagagatcctgaatgttgattga	5006		
Db	61	GATTCTTTCTGATACTGTGAATAATGAAGAACACACAGAGAGATCCTGAATGTTGATTGA	120		
QY	5007	tctaattgacagaattgccaaggttaccagacctctccaga--gcgattgacctgtgt	5063		
Db	121	TCTAATGTACAGAAATTGCCAAGGGTTACCAAGCCTCTCCAGATCTCGATTGACCTGGTT	180		
QY	5064	gcagaacatgycgagcgaagcactctcagaacgaagcaatcatgctgaaagctgcacagtgtc	5123		
Db	181	GCAGAACATGGCAGGCAAGCAGCAGCTCAGAACGAACATCATGCTGAAGCTGCACAGTGTCT	240		
QY	5124	agtcacactcagcagcactgttctgtgaatatattgagcatgctgagagccggaatatct	5183		
Db	241	AGTCCACTCAGCAGCAGCAGCTGTGTGCTGAATATTGAGCATGCTGAGAGGACCCGGAATATCT	300		
QY	5184	tcctgtgagatgtgtlaacattcagaatatltcatctaattttagaagaatctgcggt	5243		
Db	301	TCTGTGGGATGTGTAAACATTTCAAGATATTTCATCTAAGTTTGAAGAATCTGCGGT	360		
QY	5244	ctcagatgattgtgtatctccagatgaagaaggtatctgctctgaaaaatacttactga	5303		
Db	361	CTCAGATGATGTGTATCTCCAGATGAAGAAGGTATCTGCTCGTGAATAACTTTACTTGA	420		
QY	5304	gtcaggaactgtggtattacttgaaacaagcagctgcttccctctctatgctgcatgta	5363		
Db	421	GTCAGGACTTGTGGGATTTACTGGAACAGCAGCTGCTTCTCTATGGCTGGCATGTA	480		
QY	5364	tgaagcagttaatgaagtttacaagaagcagctgcttccctctctatgctgcatgta	5422		
Db	481	TGAAGCAGTTAATGAAGTTTACAAGTAAGTATCTCTATTCATGAAGCTAATCGGGATG	540		
QY	5423	caagaactatccacaatltcatgtgtaaaactlcaagaagcatlccagcaaaatgttcatc	5482		
Db	541	CAAGAACAATTCACACAATTCAATGTAACCTTCAAGAAGCATTCAGCAAAATGTTTCATC	600		
QY	5483	agagtactgctgagcgagcagatgtttgacactatlttcgtgtgtgtt--tlatggaac	5540		
Db	601	AGGATGTGAGC-----CGGATGTTGGCAGCTTATTTGCTGTGGGTTTCTAATGGAAC	654		
QY	5541	caagttcgg-----gatttggatgaaacaagaattgttttaacaaggagcctgcaat	5591		
Db	655	CCAAGTCCGGGGCAATTTGGGCTTGAACCAAGCAATTTGTTCCAGGGAGGCTTGCCAATA	714		
QY	5592	aaccaacttgcagagatatctcac-agattgagaggtatttaccggaagaagattlgag	5650		
Db	715	ACCCAAACTTGGCGAGATTTCTCACAAGATGGGAGGATTTACGCGAGAAGACTTGGCC	774		
QY	5651	aggatgtgttgaagtaatcaagaactctaatcctgttagacaagtgttaaatatgataccta	5710		
Db	775	AAGCATGTTGTGAAGCTATACAAAGAAGCTATCTGTACAGGTTATTAACTACAGGCTT	834		
QY	5711	acaagccatatatccagattaccta 5735			

Db 835 ATTACTCCATGTGCCAATCTCCTA 859

RESULT 11

AA429436

LOCUS

DEFINITION 588 bp mRNA linear EST 16-OCT-1997
zw47a12.r1 Soares-total_fetus_Nb2HF8_9w Homo sapiens cDNA clone
IMAGE:773182 5' similar to WP:F46H5.4 CE04590 RAT TRG GENE PRODUCT
; mRNA sequence.

AA429436

ACCESSION

VERSION AA429436.1 GI:2112516

KEYWORDS

SOURCE EST.

ORGANISM human.

Homo sapiens

REFERENCE

AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 588)
Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J., Moore,B.,
Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wylie
,T., Waterston,R. and Wilson,R.
WashU-Merck EST Project 1997

TITLE Unpublished (1997)

JOURNAL

COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -28m13 rev2 ET from Amersham
High quality sequence stop: 495.

FEATURES

source

1..588
location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:773182"
/dev_stage="8-9 weeks"
/lab_host="DH10B"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from pooled 8-9 week
(total) fetus material with a Not I - oligo(dT) primer [5'
TGTTACCAATCTGAAGTGGAGCGCGCTTAATTTTCTTTTCTTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 202 a 99 c 124 g 163 t

ORIGIN

Query Match 8.6%; Score 588; DB 9; Length 588;
Best Local Similarity 100.0%; Pred. NO. 5.2e-125;
Matches 588; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 744 accagatgaagaaaccaatagaacgcttagtgccttatatacccaagaacattt 803
|||||

Db 1 ACCAGATGAGGAAGAACCAATAGAACGCTTAGTGTTCCTGATATACCCAAGACATT 60
|||||

QY 804 tggtcnaagactctctgtanaaatgcttatcactcaagtttgaattgaattgaacccat 863
|||||

Db 61 TGGTCAAGAGACTTCTGTAAATGCTTATCCTCACTCAAGTTTGAATTTGAATTTGAACCCAT 120
|||||

QY 864 ttgtgcaagtttgcttatatagtgtcgaagaaagaaagatttcgaaaaacttcta 923
|||||

Db 121 TTTTCAAGTTTGGCTTATATGATGTCAAGGAAAAAGAAAGATTTCAGAAAACTTTTA 180
|||||

QY 924 ttgtgaccttaattcttagcagatgaagaggtgttaacgtccacatgtacacactgtgc 983
|||||

Db 181 TTTTGACCTTAATTCAGACAGATGAAGGGTTGTTACGTCCACATGTACCACTGCTGC 240

QY 984 cattactacccttgcaagatcagcaattttttctatcactatcccttcccaagatgttt 1043
|||||

Db 241 CATTTACTACCCTGGCAGATCAGCAATTTTCTATCATTATCCTTCCCAAGATGTTT 300
|||||

QY 1044 tcttgtaataaagctagaaaaaagtcctctacagcaagagacatltggagagtgtcagaacc 1103
|||||

Db 301 TCTTGTATTAAGCTTAGAAAAAGTCTCTACAGCAAGAGACATTGGAGTGTGCAGAAC 360
|||||

QY 1104 atatatgtttccaagaagcagatgcccaccaagaaltaaagaaaaacttgagaactgaa 1163
|||||

Db 361 ATATATGATTTTCAAGAGACAGATGCCACCAAGAATAAGAAAACTGGAGAAACTGAA 420
|||||

QY 1164 gagtcaagcagatcagttttgccaagacttggggaataatcgcacgtcttctgtgac 1223
|||||

Db 421 GAGTCAAGCAGATCAGTTTGGCCAAAGACTTGGGAAATATCGCATGCTTTGCTTGAC 480
|||||

QY 1224 tgcataccatttaataatattgttagcagctgtgaggtttggaagagattctacaga 1283
|||||

Db 481 TGCAATCCATTTAATGAATATTGTAGCAGTGTCTGGAGTTTGGAAGAGATTCTACAGA 540
|||||

QY 1284 agtagaatcagctactgagagaacgaaaggtcttgtgcagagaggag 1331
|||||

Db 541 AGTAGAATCAGTACTCGAGAACGAAAGGCTCTTGCTCAGAGAGAG 588
|||||

RESULT 12

BC020473/c

LOCUS

DEFINITION 1510 bp mRNA linear HTC 03-JAN-2002

ACCESSION BC020473

VERSION BC020473.1 GI:18042973

KEYWORDS HTC.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1510)
Strausberg,R.
Direct Submission
Submitted (03-JAN-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov

COMMENT Contact: MGC help desk
Email: cgaps-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: villalob@bcm.tmc.edu.
Villalon, D.K., Luna, R.A., Hale, S.M., Hulyk, S., Lu, X., Garcia,
A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W.,
Muzny, D.M., Gunaratne, P., Yoon, V., Kowis, C., Martin, R.,
Lawrence, S., Richards, S., Gibbs, R.A.

FEATURES

source

1..1510
location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3853958"
/tissue_type="Colon, adenocarcinoma"
/clone_lib="NIH_MGC_65"
/lab_host="DH10B"

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 13 Row: b Column: 12
This clone has the following problem: no polyA-tail.

Matches 699; Conservative 0; Mismatches 100; Indels 7; Gaps 6;	
QY 3872	tgataagccagaccgttgcacatggtcgaatgcaggagacatcgttcctcaactaacaagc 3931
Db 13	TCATAAGCCAGACGGTGGCCATGGCAATTGCAGGAACATCCGCTCCCTCAGCTGCACAAGAC 72
QY 3932	ctggcagtttcctcctcaagtcacagagtggtcaggcaacacactacctttcagcagaat 3991
Db 73	CTGGCAGTTTCTCCTCAGCTCAACGAGTGGCCGGCAGCACACCACTTCTCAGCAGAGT 132
QY 3992	caagtgcgaagccttttgatctgtctactttgggtctcaaaaatgcagatgaaacagttc 4051
Db 133	CAAGCAGAGGCTTTTGATCTGTCTGCTTGGGTTCTCAAGAATGCAGATGAACGGTTC 192
QY 4052	tacagaatgtgtttacagatctctcagctcttcagctaaaccggtctattagatctgttc 4111
Db 193	TGCAGAAATGGTTTACAGATCTCTCAGTCTCCAGCTGAACCCGGTTGTAGATCTGCTTT 252
QY 4112	atctctgtgtctcttgccttgccttgccttgccttgccttgccttgccttgccttgcct 4170
Db 253	ACCTGTGTGTATCTTCTTGTGAGTACAAAGGAAAAACGGTATTTGAACGAATGAATAGT 312
QY 4171	ttgaccttgaagaaatcaaaagacatgagcaaaagcttgaagaagctattcttggagc 4230
Db 313	TTAACTTTTAAGAATCAAAAGACATGAGAGCCAAAGCTTGAAGAGCCATTCTGGGGAGC 372
QY 4231	atagtgccaggtcaagaatggtacggtcgaagccgagacagctcgagagaagccatct 4290
Db 373	ATCGGTGCCAGCAGGAATGTACGGGGAGCCGAGGGCAGCTCGAGAGAGCCCATCT 432
QY 4291	ggaagtgccttctggaagtcaagaatattgaggtgaggaat-gatatgactcactggc 4349
Db 433	GGAAGCGCCTTTGGGAGCCAAAGAAACCTGCGGTGAGCAAGACATGACTCAGTGGCG 492
QY 4350	tcaaacacagagaagaacttgacaatcaagagcagagattgaaacacgaagcactgattga 4409
Db 493	TCAGAACTCAGAGAACTCGATAAGTCAAGACAGAGATAGAACACGAAGCACTGATTGA 552
QY 4410	tggaaaccttggtctacagaagaacaaacttaattcatttagatatacattagaattgtgtca 4469
Db 553	TGGAACCTTGCTACTGGAAGCAAACTCATCATCTTAGACACGCTAGAGATCATGCTCCA 612
QY 4470	gaccgtttctgttaacggaatccaaagagacatctctgtgtgagtgctaaaagtgcact 4529
Db 613	GACTGTGTCTTAACGGAATCCAAAGAGATATCTGGGTGTGCTAAAGTGTCTCT 672
QY 4530	acacagcatggtcctgtlaaaccaagtgcagtttatacacaactgtttgtctacacagag 4589
Db 673	ACAGAGCATGGCCTGCAACC-AAAGTCAGTCTATCTGACAGCACTGCTTGCCACGGCAGAG 731
QY 4590	agccttggttcaaaagtttcctgaactctatttgaag--aagagacagagcagtgctgc- 4646
Db 732	AGCCCTGGTCTC-AAAGTTCCTGAGCTCTTGTTCGAGGGAAGCAGACCGAGCAGTGTGG 790
QY 4647	tgattatgcctcaggtctctcgac 4672
Db 791	GGATCTGTGCTCCGCTTCTCCGGC 816
RESULT 14	
LOCUS AJ397707	796 bp mRNA linear EST 25-JAN-2001
DEFINITION AJ397707 dkfz426 Gallus gallus cDNA clone 30k15r1, mRNA sequence.	
ACCESSION AJ397707	
VERSION AJ397707.1	GI:7130918
KEYWORDS EST.	
SOURCE chicken.	
ORGANISM Gallus gallus	
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae;	
Phasianinae; Gallus.	
REFERENCE 1 (bases 1 to 796)	
AUTHORS Abdurakhmanov,I., Lodysyn,D., Geroth,P., Arakawa,H., Law,A., Plachy	

TITLE		JOURNAL		MEDLINE		COMMENT		FEATURES		BASE COUNT		ORIGIN	
								source		210 a 186 c 202 g 198 t			
								1..796					
								/organism="Gallus gallus"					
								/strain="CB"					
								/db_xref="taxon:9031"					
								/clone="30k15r1"					
								/clone_lib="dkfz426"					
								/tissue_type="Bursa of Fabricius"					
Query Match		8.3%;		Score 565.6;		DB 9;		Length 796;					
Best Local Similarity		82.8%;		Pred. No. 8.7e-120;									
Matches 659;		Conservative 0;		Mismatches 134;		Indels 3;		Gaps 1;					
QY	4465	gttcagaccglttctgttaacggaatccaaagagagcatctctgtgtgagtgctaaagt										4524	
Db	1	GTACAGACTGTTCTGTGACGGAGTCCAAAGAGAGTATCTCGCGGGGCTGCTGAAGGTG										60	
QY	4525	ctactacacagcatggtcctgtlaaccaaagtgcagtttatactacaacactgtttgttaca										4584	
Db	61	CTTCTGCACAGCATGGCCTGCCAACCAAGTGCATTTATCTCCAACTGCTTCCGCACA										120	
QY	4585	cagagagccttggttcaaaagtttcctgaactcttatttgaagaagagacagagcagttg										4644	
Db	121	CAGAGAGCTTTGGTCTCAAAAGTTCCCTGAGCTGCTGTTGAGGAGAGACTGAGCAGTGT										180	
QY	4645	gctgattatgcctcaggtctctccgacactgtagcagtagcatcggtacatacgttca										4704	
Db	181	GCAGATCTGTGCTGAGGCTCCTGAGGCACTGTAGCAGCAGCATGACACAATACGGTCA										240	
QY	4705	cacccagtgctcctccttacctactaataagggcaaaacttggagattggaataacttt										4764	
Db	241	CACGCGAGTGCCTCACTTACCTACTCATATGAGGCAAACTTGGAGATTGGCAATACTTC										300	
QY	4765	gccagggttaaatgycaggtaccaatgtcactatctccttgtgtggcacatctcagaat										4824	
Db	301	GCCAGAGTGAATAATGCAGGTGACTATGTCTCTCTCGTCCCTGGTGGGACATCCCAGAAAT										360	
QY	4825	ttaaataagaattcttaagaacgttctcctaaagactatatattgaacatatgtcgaagaat										4884	
Db	361	TTCAATGAGGAATTTTACGACGTTCTTGAAGACTAATTTTAACGTATGCTGAAGAAGAT										420	
QY	4885	ctggaattgagggaaacaacatttccctgaatcaggtccagagatcgtgtttcaatccat										4944	
Db	421	CTGGAACCTTAGAGAGACAACGTTTCTGATCAGGTCCAAAGATCTGGTTCACACTCCAC										480	
QY	4945	atgattcttctgaactgtgaaatgaaggaacaccagagagatcctgaatgttgatt										5004	
Db	481	ATGATCCTCTCTGACACAGTTAAATGAGGAGACACCAAGAGATCCAGAAATGCTAATT										540	
QY	5005	gatctaattgacagaattgccaaggtgtaccagacctctccaga---ggcattgacctgg										5061	
Db	541	GACCTGATGTACAGGATTGCAAAAGGGCTATCAGAAATTCCTTGACCTGCGCTTGACCTGG										600	
QY	5062	ttgcagaacatggcaggaagcactcagaacgaagcaatcatgtcgaagctgcacagtg										5121	
Db	601	TTACAGAACATGGCTGGAAGCACACTCGGAGAGAGAGCAATCATGCAGAGTACAGACAGTGT										660	
QY	5122	ctagtccactcagcagcactgttgcctgaatatattgagcattgtgagagccgggaatat										5181	
Db	661	TTGGTCCACTCTGCAGCCTTGGTGGCTGAGTATCTGAGCATGCTCGAGGACCGAAATAC										720	

OM of: US-09-737-246-2 to: EST:* out_format : pfs
Date: Jul 16, 2002 1:13 PM
About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
-MODEL=framed-p2n.model -DEV=x1h
-O=/cgn2_1/USPTO.spool/US09737246/runat_16072002_075429_26451/app-query.fasta_1.2180
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -GAPOP=12.000 -GAPEXT=4.000
-MINMATCH=0.100 -LOOPEXT=0.000 -LOOPEXT=0.000 -GAPOP=4.500
-GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
-FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
-DELEXT=7.000 -START=1 -MATRIX=biosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0
-ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -HEAPSIZE=500
-MINLEN=0 -MAXLEN=200000000 -USER=US09737246@cgn1_1_4507
-NCPUP=6 -ICPU=3 -LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30
-NO_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-09-737-246-2
Query length: 2090
Database: EST:*
Database sequences: 13736207
Database length: -1841457050
Search time (sec): 3201.680000

Sequence	Strd Orig	ZScore	EScore	Len	Documentation
gb_htc:AK016495	+	3187.00	5492.05	1.2e-296	2710 ! AK016495 Mus musculus adult ma
gb_htc:BC018599	+	2155.50	3709.13	2.4e-197	1311 ! BC018599 Mus musculus, clone I
gb_htc:AK004816	+	1821.50	3119.20	1.7e-164	2694 ! AK004816 Mus musculus adult ma
gb_htc:BC020473	-	1659.00	2844.46	3.4e-149	1510 ! BC020473 Homo sapiens, clone I
gb_htc:AK010755	+	1396.50	2388.32	8.8e-124	1507 ! AK010755 Mus musculus ES cells
gb_est1:AJ397707	+	1293.50	2217.76	2.8e-114	796 ! AJ397707 AJ397707 dkfz426 Gallu
gb_htc:AK018051	+	1219.00	2078.18	1.7e-106	1712 ! AK018051 Mus musculus adult ma
gb_est2:BG247108	+	1209.50	2069.71	4.9e-106	931 ! BG247108 602359776F1 NCI_CGAP_M
gb_htc:BC018075	+	1199.00	2042.23	1.7e-104	1873 ! BC018075 Homo sapiens, clone I
gb_htc:BC018076	+	1199.00	2042.23	1.7e-104	1873 ! BC018076 Homo sapiens, clone I
gb_est2:BI689126	+	1193.00	2041.95	1.7e-104	869 ! BI689126 603315930F1 NCI_CGAP_M
gb_est2:BG872239	+	1146.00	1962.31	4.7e-100	745 ! BG872239 602790848F1 NCI_CGAP_M
gb_htc:AK013336	+	1106.00	1894.89	2.7e-96	636 ! AK013336 Mus musculus 10, 11 da
gb_est2:BI770146	+	1104.00	1888.45	6.1e-96	796 ! BI770146 603053332F1 NIH_MGC_12
gb_est2:BF309712	+	1092.00	1866.00	1.1e-94	898 ! BF309712 601891854F1 NIH_MGC_17
gb_est1:AW919177	+	1088.00	1863.61	1.5e-94	636 ! AW919177 EST350481 Rat gene ind
gb_est2:BI088758	+	1085.00	1855.45	4.2e-94	795 ! BI088758 602851547F1 NIH_MGC_10
gb_est2:BI101016	+	1076.00	1841.91	2.4e-93	678 ! BI101016 602886468F1 NCI_CGAP_M
gb_est2:BE891826	+	1072.00	1830.39	1.0e-92	958 ! BE891826 601434272F1 NIH_MGC_72
gb_est2:BF717504	+	1029.00	1760.27	8.4e-89	676 ! BF717504 fd45f10.Y1 zebrafish w
gb_est2:BF159615	+	1027.50	1753.38	2.0e-88	935 ! BF159615 601767221F1 NCI_CGAP_M
gb_est2:BI737331	+	1026.00	1753.50	2.0e-88	761 ! BI737331 603357985F1 NIH_MGC_94
gb_est2:BI820302	+	1024.50	1748.28	3.9e-88	927 ! BI820302 603036864F1 NIH_MGC_11
gb_est2:BE531136	+	1022.00	1745.80	5.4e-88	805 ! BE531136 601278315F1 NIH_MGC_39
gb_est2:BI078974	+	1001.00	1709.78	5.5e-86	777 ! BI078974 602873230F1 NCI_CGAP_M
gb_est1:AA429436	+	999.00	1709.98	5.3e-86	588 ! AA429436 zw47a12.r1 Soares_tot
gb_est1:AW742035	+	988.00	1688.72	8.1e-85	692 ! AW742035 up52c04.Y1 Soares_mous
gb_est2:BI732553	+	983.00	1680.61	2.3e-84	662 ! BI732553 603355483F1 NIH_MGC_94
gb_est2:BE571095	+	974.50	1659.39	3.5e-83	1079 ! BE571095 601334443F1 NCI_CGAP_M
gb_est2:BI159195	+	973.50	1661.97	2.5e-83	778 ! BI159195 602922704F1 NIH_CGAP_M
gb_est2:BG677746	+	968.00	1650.73	1.1e-82	884 ! BG677746 602625421F1 NCI_CGAP_M
gb_est1:AL650546	+	967.50	1653.48	7.5e-82	672 ! AL650546 AL650546 XGC-neurula_9
gb_est2:BI854143	+	959.50	1636.60	6.5e-82	842 ! BI854143 603381452F1 NIH_CGAP_M
gb_est2:BF141712	+	954.50	1629.67	1.6e-81	737 ! BF141712 601790772F1 NCI_CGAP_M
gb_est2:BF120001	+	954.00	1630.50	1.4e-81	648 ! BF120001 601758086F1 NCI_CGAP_M
gb_est2:BE857715	+	953.00	1628.64	1.8e-81	654 ! BE857715 7946g03.x1 NCI_CGAP_M
gb_est1:AA275395	+	942.00	1611.41	1.6e-80	567 ! AA275395 va92f11.r1 Soares_mous
gb_est2:BG678867	+	929.00	1583.36	6.0e-79	857 ! BG678867 602624767F1 NCI_CGAP_M
gb_est2:BM310731	+	926.00	1583.42	6.0e-79	575 ! BM310731 ig47g06.Y1 HR85_islet
gb_est2:BE698863	+	924.00	1580.77	8.4e-79	540 ! BE698863 RC4-NN0025-120600-016

gb_est2:BG080811	+	913.00	1560.38	1.1e-77	595 ! BG080811 H3057607-5 NIA Mous
gb_est2:BG927424	+	906.00	1547.67	5.9e-77	620 ! BG927424 HNC53-1-E8.R HNC (H
gb_est2:BF036757	+	903.00	1542.76	1.1e-76	606 ! BF036757 601459988F1 NIH_MGC
gb_est1:AI653716	-	896.50	1530.30	5.4e-76	662 ! AI653716 wb36e07.x1 NCI_CGAP
gb_htc:AK014226	+	884.00	1502.47	1.9e-74	1051 ! AK014226 Mus musculus 13 da

seq_name: gb_htc:AK016495

seq_documentation_block:

LOCUS AK016495 2710 bp mRNA linear HTC 19-JAN-2002
DEFINITION Mus musculus adult male testis cDNA, RIKEN full-length enriched library, clone:4931431C02:homolog to KIAA1395 PROTEIN (FRAGMENT), full insert sequence.
ACCESSION AK016495
VERSION AK016495.1 GI:12855259
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (strain:C57BL/6J) adult male testis cDNA to mRNA, clone_lib:RIKEN full-length enriched mouse cDNA library
clone:4931431C02.

ORGANISM

Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 (sites)
Carninci,P. and Hayashizaki,Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
10349636

REFERENCE

2 (sites)
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159

REFERENCE

3 (sites)
Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861

REFERENCE

4 (sites)
The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5 (bases 1 to 2710)
Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A., Arakawa,T., Baldarelli,R., Bono,H., Brownstein,M., Bult,C., Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hill,D., Hiramoto,K., Hiraoka,T., Hori,F., Hume,D., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Quackenbush,J., Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Schriml,L., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yamanka,I., Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.

REFERENCE

Direct Submission
Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC),

TITLE

JOURNAL
JOURNAL

TITLE

JOURNAL
JOURNAL

TITLE

JOURNAL
JOURNAL

TITLE

JOURNAL
JOURNAL

TITLE

JOURNAL
JOURNAL

TITLE

JOURNAL
JOURNAL

TITLE

JOURNAL
JOURNAL

TITLE

JOURNAL
JOURNAL

1609	rgSerLeuLysThrIleLeuThrTyrAlaGluGluAspLeuGluLeuArg	1625
922	GTGCTCCCTGTGGGGACAACCTCAGAACTTTAGTGAAGACATTGGAGAA	971
972	AGTCCCTCAGACCACTTCGTACCTACGCAGAGGAGACATAGGGCTGAGG	1021
1626	GluThrThrPheProAspGlnValGlnAspLeuValPheAsnLeuHisMe	1642
1022	GACAGCACCTTCGCTGAGCAGGTCACAGACCTCATGTTCACCTGCACAT	1071
1642	IleLeuSerAspThrValLysMetLysGluHisGlnGluAspProGluM	1659
1072	GATCCTGACAGACAGCGTGAAGATGAAGGAACACAGAGACCCTGAGA	1121
1659	etLeuIleAspLeuMetTyrArgIleAlaLysGlyTyrGlnThrSerPro	1675
1122	TGCTCATGAGACCTCATCTACAGAAATTGCTCGGGATACCAAGCCTCCCA	1171
1676	Glu...ArgLeuThrTrpLeuGlnAsnMetAlaGlyLysHisSerGluAr	1691
1172	GACCTGCGGCTGACATGGCTGCAGAACATGGCTGGCAACATGCAGAGCT	1221
1691	gSerAsnHisAlaGluAlaAlaGlnCysLeuValHisSerAlaAlaLeuV	1708
1222	GGGCAATCATGCAGAGGCTGCCCAAGTGCATGGTCCACGCCCTGCCCTGG	1271
1708	AlaAlaGluTyrLeuSerMetLeuGluAspArgLysTyrLeuProValGly	1724
1272	TGGCCGAATACCTTGCCCTCTCGAGGACAGTCCGCACTGCTGTGGGC	1321
1725	CysValThrPheGlnAsnIleSerSerAsnValLeuGluGluSerAlaVal	1741
1322	TGCTGTCTCTCCAGAACTCTCATCTCAACAGTGTGAGAGAGTCTGCCAT	1371
1741	IserAspAspValValSerProAspGluGluGlyIleCysSerGlyLysT	1758
1372	CTCTGATGACATCTGTGCGCAGATGAGAGGGCTTCTCTCTAGGAAGA	1421
1758	yrPheThrGluSerGlyLeuValGlyLeuLeuGluGlnAlaAlaLaser	1774
1422	ACTTCACAGAACTGGGTCTGGTGGGGCTGCTGAGCAGGGCGGCTAC	1471
1775	PheSerMetAlaGlyMetTyrGluAlaValAlaSnGluValTyrLysValLe	1791
1472	TTCAACCATGGGTGTCTGTATGAAGCGGTGAACGAAGTCTACAAAAACCT	1521
1791	IleProIleHisGluAlaAsnArgAspAlaLysLysLeuSerThrIleH	1808
1522	TATCCCATCTCTGAAGCCACAGAGACTACAGAAGCTGGCTGCGGTGC	1571
1808	IscGlyLysLeuGlnGluAlaPheSerLysIleValHisGlnSerThrGly	1824
1572	ACGGGAACCTGCAGAGAGGCTTTCACCAAGATTATGCACAGAGCTCTGGC	1621
1825	TrpGluArgMetPheGlyThrTyrPheArgValGlyPheTyrGlyThrLy	1841
1622	TGGGAGCGTGTATTTGGGACATATTCCGAGTGGGCTTCTATGGCACACG	1671
1841	sPheGlyAspLeuAspGluGlnIuPheValTyrLysGluProAlaIleT	1858
1672	ATTGGTGACCTGGATGAACAAGAGTTTGTACAAAGGAACCGTCAATCA	1721
1858	hrLysLeuAlaGluIleSerHisArgLeuGluGlyPheTyrGlyGluArg	1874
1722	CGAAGCTTGACAGAGATCTCACACCGGCTGAGAGAGTTCTATACGGAAGG	1771
1875	PheGlyGluAspValValGluValIleLysAspSerAsnProValAspLy	1891
1772	TTGCGGAGATGATGTGTAGAGATCATCAAAAGATTCTAACCCAGTGGAACA	1821
1891	sCysLysLeuAspProAsnLysAlaTyrIleGlnIleThrTyrValGluP	1908

1822	GTCCAAAGCTGGACCCACAGAAAGCGGTACATACAGATTAACCTATGTGGAGC	1871
1908	roTYrPheaspThrTYrGlumetLysaspArgileThrTYrPheaspLys	1924
1872	CACATTTTCGACACTTATGAGCTCAAGGATCGGGTGAGCCTTACTTCGATCCG	1921
1925	AsnTYrAsnLeuArgargPheMetTYrCysThrProPheThrLeuAspGl	1941
1922	AACTATGGGCTGCGGGCCCTTCCTCTTGACACACCCTTCACACCAGATGG	1971
1941	YArgAlaHisGlyGluLeuHisGluGlnPheLysArgLysThrIleLeuT	1958
1972	ACGTGCGCACGAGAGATTGGCCGACACAGACAACAACGCAAGACGCTGCTGA	2021
1958	hrThrSerHisAlaPheProTYrIleLysThrArgValAsnValThrHis	1974
2022	GCACGGAGCATGCGCTTCCCTACATCAAGACACGCATCCGAGTGTGCCAC	2071
1975	LysGluGluIleIleLeuThrProIleGluValAlaIleGluaspMetGl	1991
2072	CGTGACGACAGACAGTGTCTGACACACCAGTGGAGGTGCCATTGAGGACATGCA	2121
1991	nLysLysThrGlnGluLeuAlaPheAlaThrHisGlnaspProAlaasp	2008
2122	GAAGAAGACCCGGGAGCTGGCCCTTGGCCACCAGGACGAGACCCCTCCAGATG	2171
2008	roLysMetLeuGlnMetValLeuGlnGlySerValGlyThrThrValAsn	2024
2172	CCAAGATGCTGCAGATGGTTCTCCAGGGGTCTGTGGACCACACTGTGAAC	2221
2025	GlnGlyProLeuGluValAlaGlnValPheLeuSerGluIleProSerAs	2041
2222	CAGGGTCCCTTGGAAAGTGGCCACAGTGTTTTGTGCAGAGATCCCAAGA	2271
2041	pProLysLeuPheArgHisHisAsnLysLeuArgLeuCysPheLysasp	2058
2272	TCCCAAGTCTTCCGACATCACACAACAACTCCGGCTCTGCTTCAGGATTT	2321
2058	heThrLysArgCysGluaspAlaLeuArgLysAsnLysSerLeuIleGly	2074
2322	TCTGCAAAAGATGCGAGGATGCACCTGAGAAAGAACAAGGCCCTGATTGCC	2371
2075	ProValGlnLysGluTYrGlnArgGluLeu	2084
2372	CCAGACCGAAGAGATACCAACCGGGAGCTG	2401

seq_name: gjb_htc:BC018599

```
seq_documentation_block:
```

LOCUS	BC018599	1311 bp	mRNA	linear	HTC 03-DEC-2001
DEFINITION	Mus musculus, clone IMAGE:3596745,		mRNA.		

VERSION BC018599.1 GI:17235789

SOURCE house mouse.

ORGANISM Mus musculus

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 1311)

TITLE Direct Submission

Gene Collection (MGC), Cancer Genomics Office, National Cancer

USA

COMMENT contact: MGC help desk

Email: cgapbs-r@mail.nih.gov

CDNA Library Preparation: Life Technologies, Inc.

DNA Sequencing by: Baylor College of Medicine Human Genome

Sequencing Center

seq_name: gjb_hlc:AK004816

seq_documentation_block:

LOCUS	AK004816	2694 bp	mRNA	linear	HTC 19-JAN-2002
DEFINITION	Mus musculus adult male lung cDNA, RIKEN full-length enriched library, clone:1200017A24:homolog to BA165F24.1.1 (NOVEL PROTEIN SIMILAR TO DROSOPHILA CG6630 AND CG11376, KIAA1058, RAT TRG (ISOFORM 1)) (FRAGMENT), full insert sequence.				

ACCESSION	AK004816
VERSION	AK004816.1
KEYWORDS	GI:12836282
SOURCE	HTC; CAP trapper.
	Mus musculus (strain:C57BL/6J) adult male lung CDNA to mRNA

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE	1 (sites)
AUTHORS	Carninci, P. and Hayashizaki, Y.
TITLE	High-efficiency full-length cDNA cloning
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)

REFERENCE	2 (sites)
AUTHORS	
Carninci, P.,	Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,

TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)

REFERENCE
3 (sites)
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,

TITLE	JOURNAL
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer	Genome Res. 10 (11), 1757-1771 (2000)

REFERENCE 4 (sites)
AUTHORS The RIKEN Genome Exploration Research Phase II Team and the
PANTOM Consortium.

TITLE	Functional annotation of a full-length mouse cDNA collection
JOURNAL	Nature 409, 685-690 (2001)
REFERENCE	5 (bases 1 to 2694)

AUTHORS

Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A., Arawaka, T., Baldarelli, R., Bono, H., Brownstein, M., Bult, C., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hill, D., Hiramoto, K., Hiraoka, T., Hori, F., Hume, D., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Quackenbush, J., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Schriml, L., Shibata, K., Shibata, Y., Shimagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, Y., Tejima, T., Toya, T., Yamamura, T., Yamanaka, I., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.

TITLE Direct Submission
JOURNAL Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Gemomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-frees@sc.riken.go.jp, URL: <http://genome.gsc.riken.go.jp/>, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT Please visit our web site (<http://genome.gsc.riken.go.jp/>) for

further details.

cDNA library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.
 Division of Experimental Animal Research in Riken contributed to
 prepare mouse tissues. First strand cDNA was primed with a primer
 [5' GAGAGAGAGAGCGCGCCGCAACTCGAGTTT TTT TTT TTT TTTVN 3'], cDNA was
 prepared by using trehalose thermo-activated reverse transcriptase
 and subsequently enriched for full-length by cap-trapper. Second
 strand cDNA was prepared with the primer adapter of sequence[5'
 GAGAGAGAGAGCATCCAGAGCTCATTTAATTTAATTAACCCCCCCC 3']. cDNA was
 cleaved with XhoI and SstI. Cloning sites, 5' end: SstI; 3' end:
 XhoI. Host: SOLR.

FEATURES	Location/Qualifiers
source	1. .2694

```

source
1. .2694
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="MGD:MGI:1907255"
/db_xref="taxon:10090"
/clone="1200017A24"
/sex="male"
/tissue_type="lung"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
22. .1563
CDS

```

polyA_signal	2677.	2682			
	/note="	putative"			
polyA_site	2694				
	/note="	putative"			
BASE COUNT	752	a	646	c	682
ORIGIN					g
					614
					t

[illegible]

```
alignment_block:
US-09-737-246-2 x AK004816 .
```

Align seg 1/1 to: AK004816 from: 1 to: 2694

1617	TyrAlaGlUGluAspLeuGluLeuArgGluThrThrPheProaspGlnVa	1633
:::	:::::	
1	TTTCCAGAAAGAGACACGCCCATGCAGACAACCTCTTTCCCATGCAGGT	50
1633	IGlnAspLeuValPheasnLeuHisMetIleLeuSerAspThrValIysm	1650
:::::	:::::	
51	GGAGGAACCTTCTCTGCATCTGCAACACCATTTCTGTACGACACAGTGAGA	100
1650	etlYsgLIuHisGlnGluAspProGluMetLeuIleAspLeuMetYrarg	1666
:::::	:::::	
101	TGAGGGAATTCACGAGAACCCCTGAGATGCTTATGGACCTCATGTACAGA	150

1667 IleAlaLysGlyTyrGlnThrSerProGlu...ArgLeuThrTripleuG1 1682
|||||:::|||||:::|||||:::|||||
151 ATTGCCAAGAGCTACAGGATCGCCTGACCTGGCGCTGACTGGCTCCA 200
1682 nasMetAlaGlySHisSerGluArgSerAsnHisAlaGluAlaIag 1699
|||||:::|||||:::|||||:::|||||
201 GAACATGGCAGAGAACAACACTAGAGAAGTCTTCACAGAGGCCGCCA 250
1699 IncysLeuValHisSerAlaAlaLeuValaGluTyrLeuSerMetLeu 1715
:::|||||:::|||||:::|||||:::|||||
251 TGTGCTGTGCTGATGCAGCCGCCCTGGTGGCCGAGTACCTGAGCATGCTG 300
1716 GluAspArgLysTyrLeuProValGlyCysValThrPheGlnAsnIleSe 1732
|||||:::|||||:::|||||:::|||||
301 GAGGACCACAGCTACCTGCCGGTGGCCAGCGTCACTTTCAGAATATTTC 350
1732 rSerAsnValLeuGluGluSerAlaValSerAspSpValValSerProA 1749
|||||:::|||||:::|||||:::|||||
351 TTCCAATGTGCTTGAGGAGTCTGCAGTCTGTGATGACACCTTGTCACTG 400
1749 spGluGluGlyIleCysSerGlyLysTyrPheThrGluSerGlyLeuVal 1765
|||||:::|||||:::|||||:::|||||
401 ATGAGGACGGCGTATGCTCTGTGCTGACTTCACTGAGAGTGGCCTGGTG 450
1766 GlyLeuLeuGluGlnAlaAlaAlaSerPheSerMetAlaGlyMetTyrG1 1782
|||||:::|||||:::|||||:::|||||
451 GGCTCTCTGGAGCAGGCTGGCGAGCTCTCAGCACGGGAGGCTTGTACGA 500
1782 uAlaValAsnGluValTyrLysValLeuIleProIleHisGluAlaAsnA 1799
:::|||||:::|||||:::|||||:::|||||
501 GACGGTTAATGAAGTCTACAAGCTGGTCAATCCCTATCTGGAGGCACACA 550
1799 rGAspAlaLysLysLeuSerThrIleHisGlyLysLeuGlnGluAlaPhe 1815
|||||:::|||||:::|||||:::|||||
551 GAGATTTCGGGAAGCTGACCTCCACTCAGACAAGCTGCAGAGGCCCTTC 600
1816 SerLysIleValHisGlnSerThrGlyTyrGluArgMetPheGlyThrTy 1832
:::|||||:::|||||:::|||||:::|||||
601 GATAACATCATCAACAAGGAC....CATAAAGAGATGTTGGGACCTA 644
1832 rPheArgValGlyPheTyrGlyThrLysPheGlyAspLeuAspGluGlnG 1849
|||||:::|||||:::|||||:::|||||
645 CTTCGAGTGGTCTTCTACGATCCCGATTTGGGATTTGGATGAGCAGG 694
1849 LuPheValTyrLysGluProAlaIleThrLysLeuAlaGluIleSerHis 1865
|||||:::|||||:::|||||:::|||||
695 AGTTCGTGTACAGAAGAACCCGCAATCAAGAGCTCCCGGAGATCTCACAT 744
1866 ArgLeuGluGlyPheTyrGlyGluArgPheGlyGluAspValValGluVa 1882
|||||:::|||||:::|||||:::|||||
745 AGACTAGAGGAGATTTATGGCCAGTGTTCGGTGCAGAGTGTGGGAAGT 794
1882 IleLeuAspSerAsnProValAspLysCysLysLeuAspProAsnLysA 1899
|||||:::|||||:::|||||:::|||||
795 GATAAAAGACTTACTCTCAGTGGACAACAACCAAGTTGATCCTAAACAAG 844
1899 IatyrIleGlnIleThrTyrValGluProTyrPheAspThrTyrGluMet 1915
|||||:::|||||:::|||||:::|||||
845 CCTACATTCAGATCACTTTGTGGAGCCTTACTTTGATGAATATGAGATG 894
1916 LysAspArgIleThrTyrPheAspLysAsnTyrAsnLeuArgArgPheMe 1932
|||||:::|||||:::|||||:::|||||
895 AAAGACCGGGTGACCTACTCTGAGAAGAATTTCACCTCCGAGGTTTCAT 944
1932 tTyrCysThrProPheThrLeuAspGlyArgAlaHisGlyGluLeuHisG 1949
|||||:::|||||:::|||||:::|||||
945 GTACACACACCCCTTCACCCCTGGAGGGGAGACCCCGGGGCGAGCTTCATG 994
1949 LuGlnPheLysArgLysThrIleLeuThrThrSerHisAlaPheProTyr 1965
|||||:::|||||:::|||||:::|||||
995 AGCAACACCGCAGAAACACCGTGTCTCACCAACATGACAGCCTTCCCTTAC 1044
1966 IleLysThrArgValAsnValThrHisLysGluGluIleIleLeuThrPr 1982

|||||:::|||||:::|||||:::|||||
1045 ATCAAGACCAGGATCCGAGTCAAGCCAGAAAGAGAGTTCGTTTGACTGCC 1094
1982 olleGluValAlaIleGluAspMetGlnLysLysThrGlnGluLeuAlaP 1999
|||||:::|||||:::|||||:::|||||
1095 GATTGAGGTTGCCATGTGAGATATGAGAAGAAGACCCCTGCAGTTAGCCG 1144
1999 heAlaThrHisGlnAspProAlaAspProLysMetLeuGlnMetValLeu 2015
|||||:::|||||:::|||||:::|||||
1145 TGGCCACTCACCAGAGACCCCTGTATGCAAAGATGCTGCAAAATGTTACTG 1194
2016 GlnGlySerValGlyThrThrValAsnGlnGlyProLeuGluValAlaG1 2032
|||||:::|||||:::|||||:::|||||
1195 CAGGGCTCTGTAGAGCCACTGTAATCAGGAGCACCTGGAGGTGCCCA 1244
2032 nValPheLeuSerGluIleProSerAspProLysLeuPheArgHisAla 2049
|||||:::|||||:::|||||:::|||||
1245 AGTGTCTTGGCTGAATTCACAGCTGACCCAAAGCTCTTACCGACATCACA 1294
2049 snLysLeuArgLeuCysPheLysAspPheThrLysArgCysGluAspAla 2065
|||||:::|||||:::|||||:::|||||
1295 ACAAGCTGAGGTGTGCTTCAAGAGAGTTCAATATGCGATGCGGAGAGGCC 1344
2066 LeuArgLysAsnLysSerLeuIleGlyProValGlnLysGluTyrGlnAr 2082
:::|||||:::|||||:::|||||:::|||||
1345 GTGGAGAAGAAGACAGCCACTCATCACCAGACAGCGGAGTACCAGCA 1394
2082 gGluLeuGlyLys 2086
:::|||||:::|||||
1395 GGAGCTGAAGAAG 1407

seq_name: gb_htc:BC020473

seq_documentation_block: 1510 bp mRNA 1linear HTC 03-JAN-2002
LOCUS BC020473
DEFINITION Homo sapiens, clone IMAGE:3853958, mRNA.
ACCESSION BC020473
VERSION BC020473.1 GI:18042973
KEYWORDS HTC.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 1510)
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (03-JAN-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgabs-r@mail.nih.gov
Tissue Procurement: AFCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: villalonebcm.tmc.edu.
Villalon, D.K., Luna, R.A., Hale, S.M., Hulyk, S., Lu, X., Garcia,
A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W.,
Muzny,D.M., Gunaratne, P., Yoon, V., Kowis, C., Martin, R.,
Lawrence, S., Richards, S., Gibbs,R.A.

FEATURES
source
1. 1510
/organism="Homo sapiens"
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov
Series: IRAK Plate: 13 Row: b Column: 12
This clone has the following problem: no polyA-tail.

```

/db_xref="taxon:9606"
/clone="IMAGE:3853958"
/tissue_type="Colon, adenocarcinoma"
/clone_lib="NIH_MGC_65"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
BASE COUNT      287 a      431 c      456 g      336 t
ORIGIN

alignment_scores:
    Quality: 1659.00      Length:      381
    Ratio: 4.621      Gaps: 0
    Percent Similarity: 94.226      Percent Identity: 81.102

alignment_block:
US-09-737-246-2 x BC020473/rev ..

Align seg 1/1 to reverse of: BC020473 from: 1 to: 1510

1704 SerAlaAlaLeuValAlaGluTyrLeuSerMetLeuGluAspArgLysTyr 1720
      ::::::::::::::::::::::::::::::::::::::::::::::::::::
1510 GCGGCCGCCCTCGTGGCTGAGTACCTCGCCCTGCTCGAGGACCACCGCCA 1461
1720 rLeuProValGlyCysValThrPheGlnAsnIleSerSerAsnValLeuG 1737
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
1460 CCGGCCCGTGGCGGTCTTCTCCAGAACATCTCATCCAAATGTGCTAG 1411
1737 LuGluSerAlaValSerAspAspValValSerProAspGluGluGlyIle 1753
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
1410 AGGAGTCCGCCATCTCGAGACATCTGTGCCCGGACGAGAGGCGCTTC 1361
1754 CysSerGlyLysTyrPheThrGluSerGlyLeuValGlyLeuLeuGluG 1770
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
1360 TGCTCCGGGAGACACTTCAGCTGAGCTGGGGCTGGTAGGGTTGCTGAACA 1311
1770 nAlaAlaAlaSerPheSerMetAlaGlyMetTyrGluAlaValAsnGluV 1787
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
1310 GGCAGCCCGGCTACTTCAACATGGCGGGCTCTACGAGGCGGTGAATGAGG 1261
1787 alTyrLysValLeuIleProIleHisGluAlaAsnArgAspAlaLysLys 1803
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
1260 TCTACAAGAACCTCATCCCATCTCTGGAAGCCCAACCGTACTACAAGAAG 1211
1804 LeuSerThrIleHisGlyLysLeuGlnGluAlaPheSerLysIleValHi 1820
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
1210 CTGGCCCGCGGTGCACGCGCAAACTGCAGAGGCGCTTCACCAAGATCATGCA 1161
1820 sGlnSerThrGlyTyrPgluArgMetPheGlyThrTyrPheArgValGlyP 1837
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
1160 CCAGAGTTCGGGCTGGGAGCGCGGTGTCGGGAGCTATTCCGCGTGGGCT 1111
1837 heTyrGlyThrLysPheGlyAspLeuAspGluGlnGluPheValTyrLys 1853
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
1110 TCTACGGGCGCCACTTCGCTGCTGCTGATGAGCAGAGATTGTGTACAAG 1061
1854 GluProAlaIleThrLysLeuAlaGluIleSerHisArgLeuGluGlyPh 1870
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
1060 GAGCCATCGATCAGCAAGCTGGCAGAGATCTCACACCGGCTGGAGAGATT 1011
1870 eTyrGlyGluArgPheGlyGluAspValValGluValIleLysAspSera 1887
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
1010 CTACACGAGAGATTGGCGACGACGCTGTTGAGATTATCAAAAGACTCTTA 961
1887 snProValAspLysCysLysLeuAspProAsnLysAlaTyrIleGlnIle 1903
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
960 ACCCTGTGGACAAGTCAAGCTTGACTCACAAAGCCCTACATCCAGATTC 911
1904 ThrTyrValGluProTyrPheAspThrTyrGluMetLysAspArgIleTh 1920
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
910 ACGTATGTGAACCGTACTTGTATACCTACGAGCTCAAGACCGGGGTGAC 861
1920 rTyrPheAspLysAsnTyrAsnLeuArgArgPheMetTyrCysThrProp 1937
```

```

|||||.....|
860 CTACTTTGACCGCAACTATGGCTTCGCACATTCCTGTTCGCACGCCGT 811
1937 heThrLeuAspGlyArgAlaHisGlyGluLeuHisGluGlnPheLysArg 1953
      ||||| ..|||
810 TCACGCCGGATGGCGCCGACACGGGGAGCTGCCCGGACACACAAGCGT 761
1954 LysThrIleLeuThrThrSerHisAlaPheProTyrIleLysThrArgVa 1970
      |||||.....|
760 AAGACGCTGCTCAGACCCGACCGACCGCTTCCCTACATCAAGACTCGCAT 711
1970 lAsnValThrHisLysGluGluIleIleLeuThrProIleGluValAlaI 1987
      ::::||| ||:::||||| ::::|
710 CCGTGTGTGCCACCGGAGGAGACGGTGTCTGACGCCAGTGAAGTGCCA 661
1987 lGluAspMetGlnLysLysThrGlnGluLeuAlaPheAlaThrHisGln 2003
      |||||.....|
660 TCGAGGACATGCAGAAAGACACGGGAGCTGGCCTTGGCCACCGAGCAG 611
2004 AspProAlaAspProlLysMetLeuGlnMetValLeuGlnGlySerValG 2020
      ||||| ||| ..|||
610 GACCCACAGATGCTAAGATGCTACAGATGTGCTTCAGGGCTCTGTAGG 561
2020 yThrThrValAsnGlnGlyProLeuGluValAlaGlnValPheLeuSerg 2037
      | ..|||
560 GCCCACCCTGAACCAAGGCTCCCTGGAAGGTGGCCCAAGTGTTTTAGCAG 511
2037 LuIleProSerAspProlLysLeuPheArgHisHisAsnLysLeuArgLeu 2053
      |||||.....|
510 AGATCCCGGAAGACCCCAAGCTCTTCCGGCATCACACAATTGCGGCTC 461
2054 CysPheLysAspPheThrLysArgCysGluAspAlaLeuArgLysAsnLy 2070
      |||||.....|
460 TGCTTCAAGACTTCTGCAAGAAATGTGAGATGCGCTGCGGAAAAATAA 411
2070 sSerLeuIleGlyProValGlnLysGluTyrGlnArgGluLeu 2084
      |:::|
410 GGCCCTGATTGGCCCGGACGACGAGAAGAGTACCACCGTGAGCTG 368

seq_name: gb_htc:AK010755

seq_documentation_block:
LOCUS AK010755 1507 bp mRNA linear HTC 19-JAN-2002
DEFINITION Mus musculus ES cells cDNA, RIKEN full-length enriched library,
            clone:2410095B20:homolog to KIAA1395 PROTEIN (FRAGMENT), full
            insert sequence.
ACCESSION AK010755 GI:12846421
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (strain:C57BL/6J) ES cells cDNA to mRNA,
        clone_lib:RIKEN full-length enriched mouse cDNA library
        clone:2410095B20.
ORGANISM Mus musculus
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (sites)
AUTHORS Carninci,P. and Hayashizaki,Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE 2 (sites)
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
          Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
          Normalization and subtraction of cap-trapper-selected cDNAs to
          prepare full-length cDNA libraries for rapid discovery of new genes
          Genome Res. 10 (10), 1617-1630 (2000)
JOURNAL 20499374
MEDLINE 11042159
PUBMED 11042159
REFERENCE 3 (sites)
AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
          Konno,H., Akiyama,J., Nishi,K., Kitsuunai,T., Tashiro,H., Itoh,M.,
          Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
```


Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujitake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

TITLE
RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)

JOURNAL
MEDLINE
PUBMED

REFERENCE
AUTHORS

TITLE	JOURNAL	REFERENCE	AUTHORS
1. The Effect of Temperature on the Rate of Reaction of Hydrogen Peroxide with Potassium Iodide	Journal of Chemical Education	1965	Smith, J. D.
2. The Effect of Concentration on the Rate of Reaction of Hydrogen Peroxide with Potassium Iodide	Journal of Chemical Education	1965	Smith, J. D.
3. The Effect of Catalyst on the Rate of Reaction of Hydrogen Peroxide with Potassium Iodide	Journal of Chemical Education	1965	Smith, J. D.
4. The Effect of pH on the Rate of Reaction of Hydrogen Peroxide with Potassium Iodide	Journal of Chemical Education	1965	Smith, J. D.
5. The Effect of Surface Area on the Rate of Reaction of Hydrogen Peroxide with Potassium Iodide	Journal of Chemical Education	1965	Smith, J. D.

TITLE
JOURNAL

COMMENT

FEATURES
source

CDS

/translation="MEQELRASLALRASPEPLVAFSHLVIDKLVRVLRPPITCGQ
MYNLGRGAFEMAHYASLVHRNLEAVQDSRGHCPPLASYAHYAFRLPGDLSLDEAF
PATVQATLARGSGRPSLYLARSKSTSSNPDLAVPGSVDEYSRILASKALDRNS
SRASSYLEASSALPQSRHTVQKLHELALQVWVSGSAVRELVLQHAFFQDLMW

	KSMELHLLGQRDTPRNVRFPGRFLLDDIALVASVGLVITRVHKTGCLCVTLGCLA L"			
BASE COUNT	269 a	490 c	413 g	335 t
ORIGIN				

```
alignment_scores:
```

Quality:	1396.50	Length:	500
Ratio:	3.474	Gaps:	4
Percent Similarity:	80.400	Percent Identity:	53.200

```
alignment_block:
```

US-09-737-246-2 x AK010755 . .

Align seg 1/1 to: AK010755 from: 1 to: 1507

567 TyrProGlnSerLeuAsnPhenAlaAsnArgGlnGlySerAlaArgAsnI1 583

3 TACCCACACAGCCTCAATTCAGCAGTCGCCAGGGCTCTGTGGCAACCT

583 eThrValLysValGlnPheMetTyrGlyGluAspProSerAsnAlaMetP 600

53 GGCTGTGGGATCCAGTACATGGCCGGTGAAGACCAGAGCCAGGCTTGC 102

600 rovalIIepheglyLysSerSerCysSerGluPheSerLysGluAlaTyr 616

103 CGGTATCTTTGGGAATCTAGCTGCAGCGAATCACCCGAGAGCCCTC 152

617 ThrAlaValValTyrHisAsnArgSerProAspPheHisGluGluLeu 633

153 ACACCAGTGGTCTATCACAACAAGTCTCCTGAATTCTACGAGGAATTCAA 202

633 svallysLeuProAlaThrLeuThrAspHisHisLeuLeuPheThr 650

203 GCTACGACTTCCTGCGCTGCGTGACCGAGAACCATCACCCTTTTACCT 252

650 hetYrHisValSerCysGlnGlnIleLysGlnAsnThrProLeuGluThrPro 666

253 TCTACCATGTCAGCTGCCAGCCCCGGCCAGGAACAGCCTTGGAGACGCC 302

667 valglytyrthrtrpIleprometleuglnasn glyargleuylsthrG1 683

303 GTGGGCTTCACTTGGATCCCTCTGTACACATGGCCGCTGAGGACTGG 352

683 yglNphecysleuprovalserleuglulysproproglinalatyrser 700

353 TCCCTTCTGCCTGCCCGTGTCCGTGGACCAAGCCTCCACCAAGCTACTCGG 402

```

700  aLeuSerProGluValProLeuProGlyMetLysTrpValAspAsnHis  716

```

403 TCCTGACCCCGAGATGTAGCGCTGCCCTGGCATGCCGCTGGGTGGACGGCCAC 452

717 LysGlyValPheAsnValGluValValAlaValSerSerIleHisThrG1 733

453 AAGGCCGTGTTCAAGTGTGGAGCTCACCGCCGTGTCGTCCGTGCACCCGCA 502

733 nasprProTyrLeuasplysPhehealaleuValasnalaleuasprGluH 750

503 GGACCCCACTTGGATAAGTTCTTCACGCTGGTACATGTTCTAGAGGAAG 552

750 isleupheprovalargileglyaspmetargilemetgluasnasleu 766

553 GGATCTTCCGTTCCGACTCAAGGAGACAGTGTGAGCTAAGGCACTATG 602

767 GIUASNGLULeuLysSerIleSerAlaLeuASnSerSerGlnLeuGI 783

603 GAGCAGGAGTTGCCGAGCCAGCCTGGCAGCCCTGGCGCTTGCCAGCCAGA 652

783 uProva1ValArgpHeuHisLeuLeuAspLysLeuIleLeuLeu 800

653 GCCCCTAGTAGCCTTTTCCACCTCCTGCTAGACAAGCTTGTCCGCTGG 702

800 allleargProovalllealaglycIntllevalAsnleuglycInala 816


```

703 TTCTGCGGCCACCCATCATTTGTGGCCAGATGTTGAACCTGGGTGAGGG 752
817 SerPheGluAlaMetAlaSerIleIleAsnArgLeuHisLysAsnLeuG 833
753 GCCTTTGAAGCTATGGCTCATGTAGCCAGCCTTGTGCACCGAAGCTGGA 802
833 uGlyAsnHisAspGlnHisGlyArgAsnSerLeuLeuAlaSerTyrIleH 850
803 GGGTGTCCAAAGACTCCCGTGGGCCACTGCCCACTGCTGGCTCTATGTCC 852
850 IStyValPheArgLeuProAsn.....ThrTyrProAsnSerSer 863
853 ACTACGCTTTCCGCTCCCTGGTGGTGTACCTCAGCCTGCAGCGGAAGCC 902
864 SerProGlyProGlyGlyLeuGlyGlySerValHisTyrAlaThrMetAl 880
903 CCTCCA.....GCACTGTGTCAAGCTGCCACACTGGC 934
880 AArgSerAlaValArgProAlaSerLeuAsnLeuAsnArgSerArgSerL 897
935 CCGTGGCTTGTGTCGCCCGCCAGCCTGTACCTGGCACGGTCTAAGAGCA 984
897 euSerAsnSerAsnProAspIleSerGlyThrProThrSerProAspasp 913
985 TCAGCAGCAGCAACCTGACCTGGCTGTGGTCCCTGGCTCTGTGACGAT 1034
914 GluValArgSerIleIleGlySerLysGlyLeuAspArgSerAsnSerTr 930
1035 GAGGTGTCCCGCATCTCCTAGCCAGCAAGGCCATCGAACCGCACTTAGC. 1082
930 pValAsnThrGlyGlyProLysAlaAlaProTyrGlySerAsnProSerp 947
1082 ..... 1082
947 roSerAlaGluSerThrGlnAlaMetAspArgSerCysAsnArgMetSer 963
1083 .....CGGGCCTCC 1091
964 SerHisThrGlu...ThrSerSerPheLeuGlnThrLeuThrGlyArgLe 979
1092 TCTTACCTCGAAGCCTCATCTCGGCCCTGCCAGCCCTCAGTCGAGACA 1141
979 uProThrLysLysLeuPheHisGluGluLeuAlaLeuGlnTrpValValc 996
1142 CACTGTGCAGAGACTGCTGTCACAGAGAGCTGGCTGTGCAGTGGGTGTTGA 1191
996 ysSerGlySerValArgGluSerAlaLeuGlnGlnAlaTrpPhePhe 1012
1192 GCGGCAGTGGCGTTCTGTAGCTAGTCTGTGCAGCATGGCTTCTTCTTC 1241
1013 GluLeuMetValLysSerMetValHisHisLeuTyrPheAsnAspLysLe 1029
1242 CAGCTCATGGTAAAGCATGGAGCTCCATCTGCTTGTGGGCCAGCGACT 1291
1029 uGluAlaProArgLysSerArgPheProGluArgPheMetAspAspIleA 1046
1292 GGACACTCCCGCAACGTGCGCTTCCCTGGCGCTTCTGTGATGACATTG 1341
1046 IaAlaLeuValSerThrIleAlaSerAspIleValSerArgPheGlnLys 1062
1342 CCGCCCTGTGTGGCTTCGTGGGCCCTGGAAGTCATCACCGAGACTCCATAAG 1391
seq_name: gb_est1:AJ397707

```

```

REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS        Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
                Phasianinae; Gallus.
                1 (bases 1 to 796)
TITLE          Abdrakhmanov,I., Lodygin,D., Geroth,P., Arakawa,H., Law,A., Plachy
                'J., Korn,B. and Buerstedde,J.M.
                A large database of chicken bursal ESTs as a resource for the
                analysis of vertebrate gene function
JOURNAL        Genome Res. 10 (12), 2062-2069 (2000)
MEDLINE        20568495
COMMENT        Contact: Buerstedde JM
                Cellular Immunology
                Heinrich-Pette-Institute
                Martinistr. 52, 20251 Hamburg, Germany
                Email: http://genetics.hpi.uni-hamburg.de/dt40est.html.
FEATURES
    source      1..796
                /organism="Gallus gallus"
                /strain="CB"
                /db_xref="taxon:9031"
                /clone="30k15r1"
                /clone_lib="dkfz426"
                /tissue_type="Bursa of Fabricius"
BASE COUNT     210 a      186 c      202 g      198 t
ORIGIN
alignment_scores:
    Quality:    1293.50      Length:      265
    Ratio:      4.937        Gaps:      1
    Percent Similarity: 98.868      Percent Identity: 96.981
alignment_block:
    US-09-737-246-2 x AJ397707 ..
Align seg 1/1 to: AJ397707 from: 1 to: 796
1482 valGlnThrValSerValThrGluSerLysGluSerIleLeuGlyGlyVal 1498
|||||
1 GTACAGACTGTTCTGTGACGAGTCCAAAGAGAGATATCCTCGCGGGGT 50
1498 IleuLysValLeuLeuHisSerMetAlaCysAsnGlnSerAlaValTyrL 1515
|||||
51 GCTGAAGGTGCTTCTGCACAGCATGGCTGCACCAAGTGCACTTTATC 100
1515 euGlnHisCysPheAlaThrGlnArgAlaLeuValSerLysPheProGlu 1531
|||||
101 TCCACACACTGCTTCGCCACACAGAGAGCTTTGGTCTCAAGTCCCTGAG 150
1532 LeuLeuPheGluGluGluIuThrGluGlnCysAlaAspLeuCysLeuArgLe 1548
|||||
151 CTGCTGTTTGAGGAGGAGACTGAGCAGTGTGCAGATCTGTGCCTGAGGCT 200
1548 uLeuArgHisCysSerSerSerIleGlyThrIleArgSerHisProSera 1565
|||||
201 CCTGAGGCACCTGTAGCAGCAGCATCAGCACAAATACGGTTCACACGCGAGTG 250
1565 laSerLeuTyrLeuLeuMetArgGlnAsnPheGluIleGlyAsnAsnPhe 1581
|||||
251 CCTCACTTTACCTACTCATGAGGGCAAAACTTTGAGATTGGCAATTAACCTC 300
1582 AlaArgValLysMetGlnValPrometSerLeuSerSerLeuValGlyTh 1598
|||||
301 GCCAGAGTGAATAATGACAGGTACTATGTCTCTCGTCCCTGGTGGGAGC 350
1598 rSerGlnAsnPheAsnGluGluPheLeuArgArgSerLeuLysThrIleL 1615
|||||
351 ATCCAGAAATTTCAATGAGGAATCTTTACGACGTTCCCTGAAGACTATTT 400
1615 euThrTyrAlaGluGluAspLeuGluLeuArgGluThrThrPheProAsp 1631
|||||
401 TAACGTATGCTGAAGAAGATCTGGAACTTAGAGAGACACAGCTTTCCTGAT 450

```


57	CTTTATGAGCCCTGTGCGAGCCAGCTGGAGCTTTGAAGGACTCGTGATGACACA	106
67	sProLeuAlaValAlaSPSerGlyProLeuArgAspLeuIleGluPhePro	84
107	CTTAAACAGCCTTGAGATGCAAGAGCTGGCCAGAGAGCTGGGGAGCTCACC	156
84	roAspAspIleGluValValTyrSerProArgAspCysArgThrLeuVal	100
157	ATGACGACCTGCATGTGGCCTTCACACCCAAAGAAATGTAGAGCTTTGCAG	206
101	SerAlaValProGluGluSer..GluMetAspProHisValArgAspCys	116
207	CACCTCTGCGCAGAGAGAGAGTGAAGCTGATGATCCTCAGCAGAGACTG	256
116	SlleArgSerTyrThrGluAspTrpAlaIleValIleArgLysTyrHisL	133
257	TGTTCAAGACCTATATTCAGAGAGTGGCTGATTGTAAACCGGAAAAACCAAG	306
133	ysLeuGlyThrGlyPheAsnProAsnThrLeuAspLysGlnLysGluArg	149
307	AAAGTTCAGAGTTTTCGAGCTTTAAAAAGACGGGATCTCGCAGAGATT	356
150	GlnLysGlyLeuProLysGlnValPheGluSerAspGluAlaProAspG	166
357	CAGAAGACGCTTCAGAAACAGACGTTTGAGTCAGAAACCTTGAAGTCAG	406
166	y.AsnSerTyrGlnAspAspGlnAspAspLeuLysArgArgSerMetSer	182
407	TGAAC.....CGACACTCAGC	423
183	IleAspAspThrProArgGlySerTrpAlaCysSerIlePheAspLeuLys	199
424	CTGCAG.....	429
199	sAsnSerLeuProAspAlaLeuLeuProAsnLeuLeuAspArgThrPro	216
430CCTGATGAGCGGCTGGAAAAACCTGCTCCAGCTTGTGAGCG	469
216	sngLgluIleAspArgGlnAsnAspAspGlnArgLysSerAsnArgHis	232
470	CTGAGGACTTTGAGAAAGAGAGAGAGGCCCGCAAGACCAATCGGCCG	519
233	LysGluLeuPheAlaLeuHisProSerProAspGluGluGluProIleG	249
520	GCTGAGCTCTTTGCCCTCTATGCCGCCCTGGATGAGAGAGATGCTGTGA	569
249	uArgLeuSerValProAspIleProLysGlnHisPheGlyGlnArgLeuL	266
570	AATACGTCACGATGATGATGTCCTAAGGAACATCTGGGCAACAGAAATAT	619
266	euValLysCysLeuSerLeuLysPheGluIleGluIleGluProIlePhe	282
620	TGGTCAAGGTGCTGACTCTGAAGTTTGAGATTGAAATTGAACCTCTGTTT	669
283	AlaSerLeuAlaLeuTyrAspValLysGluLysLysLysIleSerGluAs	299
670	GCCAGTATTGCCCTCTATGACGTTTAAAGAAAGAAAAAGATCTCAGAAAA	719
299	nPheTyrPheAspLeuAsnSerGluGlnMetLysGlyLeuLeuArgProH	316
720	TTTCCACTGTGACCTGAACTCCGACCAAGTTCAAAAGGTTTCTCGCAGCTC	769
316	isValProProAlaAlaIleThrThrLeuAlaArgSerAlaIlePheSer	332
770	ACACACCCCTCGATTGACCCATCGAGTCAGGCTAGGTCTGCCGTGTCTCT	819
333	IleThrTyrProSerGlnAspValPheLeuValIleLysLeuGluLysVa	349
820	GTCACCTACCCATCTTCAGACATCTACCTTGGTGTCAAGATTGAAAGGT	869
349	IleuGlnGlnGlyAspIleGlyGluCysAlaGluProTyrMetIlePheL	366
870	CCTTTCAGCAAGGAGAGATTGACAGACTGTGCAGAAACCTACATGATCATCA	919

366 ysgluAlaaSpIatThrLysAsnLysGluLysLeuGIuLysLeuYsSer 382
|||||:::|||||:: |||:::|||||:::|||||
920 AAGAAAGCGATGGTGAAGAAGATTAAAGAAAAGTTGAAAAAATAACTAAACCTT 969
383 GlnlaaspGlInPheCysGlnArgleuGIlySTyrrArgetProPheal 399
|||||:::|||||:::|||||:::|||||
970 CAAGCTGAATCCCTTCGTGCCAACGTTTTGGGGAATAACCGGATGCCCTTCGC 1019
399 aTrpThralalleHisleuMetAsnileValserSeraLaGIySerleug 416
|||||:::|||||:::||| ::::: :::: :::::
1020 CTGGGCCCATTAAGCTTAGCAAGCTTCCTC.. AACATCTCACCCCTTG 1066
416 luArgasPserThrGIuValGIuileserThrglGIuArglysGIySer 432
|||||:::|||||:::||||| ::::::: |::: :::::
1067 AAAGGAGAGACACAGATGTGGAGCCTGGGGTTGGG.... AGGACTCT 1110
433 TrpserGIuArgatgAsnSerSerilevalGIyArgArgSerleuGIuAr 449
:::|||:::|||||:::||| ::||| |||||
1111 GTGGGTGAGAAAGAGAGAGCTTGTCC.. CAATCCAGGAGGCCCTCTGAGCG 1157
449 gThrThrSerGIyaspaSPalaCyasaSnleuthrSerPhearGProLat 466
||||| ||| ::::: :::: ::|||::: ::::
1158 AACCCCTCTCCTTGAGGAAAATGGAGTTGGATCCGACCTCAAAGACCACCA 1207
466 hrleuthrValThrAnpPhePhelySGlnGLyAsparGLEuSerASP 482
|||||:::|||||:::|||||:::|||||
1208 CCTGGCCACCAACAATCTTCTTCCAACAGAGAGAGATCGCCTTAGTGAT 1257
483 GluaspleuTyrlYspHeuleulaaspmetarGargProSerSerValle 499
|||||:::|||||:::|||||:::||| ::||| |||||
1258 GAAGACTGTTCAGATTTTATAGCTGACTACAGAGATCTTCATCCCTACA 1307
499 uARgArgLEuaRgProIleThraLIaglnleuLYstleasPIleserPro 516
|||||:::|||||:::||| ::::: |||::: |||||
1308 GCGAAGAGTCAAAATCCATCCACAGGCTCACTGAGGCTGGAGATATCCCCAG 1357
516 laProglUasnProHIStyrCYsleuThrProglUenleuGlnValLyS 532
|||||:::|||||:::|||||:::|||||
1358 CTCGCCGACGTGATGAACGTGCTGCTGACGCCCGAGGTGCTGCCAGTCAAA 1407
533 leuTYrPROASpSERarGValARGproTHrarGluILEugluPhePR 549
:::|||:::|||||:::|||||:::||| ::||| |||||
1408 CcTTTtccTGAAAAAtcgAgCGcGTCCAcAcAcAAgAGaGaTttTGAAAttTtcC 1457
549 calaarGasPaVAltyrVaIPROASNthrThryrArg.....A 562
| |||||:::|||||:::|||||:::|||||
1458 GaTCCGGGAGGTgtACgTCCctCaCaActGTgtACAGGTgaGAaATgtGGG 1507
562 snleuLeuTyrlletyrPrOglinSerleu 571
|||||:::|||||:::|||||
1508 CCcTtGtTtgATaATacACTgcTcTcTC 1536

seq_name: qb_est2:BG247108

seq_documentation_block:

LOCUS	EG247108	931 bp	mRNA	linear	EST 13-FEB-2001
DEFINITION	602359776F1	NCI_CGAP_Mam1	Mus musculus	CDNA clone	IMAGE:4488235 5',

mRNA sequence.

REVISION	00247108	GI:12756923
VERSION	BG247108.1	

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

Eukaryota; M

REFERENCE 1 (bases 1 to 931)

AUTHORS NIH-MGC <http://mgc.>

TITLE National Institutes

JOURNAL Unpublished (1999)

COMMENT

Contact: Robert Str...

Email: cyapbs-r@gmail.com

Tissue Procurement:

CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM10334 row: g column: 20
High quality sequence stop: 728.

FEATURES
source
location/Qualifiers
1..931

/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4488235"
/clone_lib="NCI_CGAP_Mam1"
/tissue_type="tumor, biopsy sample"
/dev_stage="10 months, virgin"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"
BASE COUNT 275 a 198 c 228 g 230 t
ORIGIN

alignment_scores:
Quality: 1209.50 Length: 272
Ratio: 4.670 Gaps: 4
Percent Similarity: 95.221 Percent Identity: 88.971

alignment_block:
us-09-737-246-2 x BG247108 ..

Align seg 1/1 to: BG247108 from: 1 to: 931

1742 SerAspAspValValSerProAspGluGluGlyIleCysSerGlyLysTy 1758
|||||
2 TCAGATGATGTGGTATCTCCAGATGAAGAAGGCATTGCTCTGGAAGTA 51
1758 rPheThrGluSerGlyLeuValGlyLeuGluGlnAlaAlaSerP 1775
|||||
52 CTTCACCTGAATCGGGACTGCTGGCTTACTGGAACAAGCGCTGCTTCC 101
1775 heSerMetAlaGlyMetTyrgluAlaValAsnGluValTyrlsValleu 1791
|||||
102 TCTCCATGCGGGGCATGTACGAAGCATTATGAGGTTTACAAAGTACTC 151
1792 IleProIleHisGluAlaAsnArgAspAlaLysLysLeuSerThrIleHi 1808
|||||
152 ATTCCATTATCATGAAGCTAATCGGGATGCAAAAGAAGCTCTCCACAATTCA 201
1808 sGlyLysLeuGlnGluAlaPheSerLysIleValHisGlnSerThrGlyT 1825
|||||
202 TGGCAAACTTCAGAGGACATTCAGCAAAATCGTCCATCAGGAT.....G 245
1825 rPgluArgMetPheGlyThrTyrrPheArgValGlyPheTyrglyThrLys 1841
:::|||||
246 GTAAGCGGATGTTGGCACCTATTTCGGTGTGGTTTATGGAACCAAG 295
1842 PheGlyAspLeuAspGluGlnPheValTyrlsGluProAlaIleTh 1858
|||||
296 TTCGGGGATTGGATGAACAGCAATTGTTTACAAGGAGCCTGCAATAAC 345
1858 rLysLeuAlaGluIleSerHisArgLeuGluGlyPheTyrglyLuarGP 1875
|||||
346 CAAACTCGCAGAGATTCTCAGAGATTGGAGGATTATGAGAAAAGAT 395
1875 heGlyGluAspValValGluValIleLysAspSerAsnProValAspLys 1891
|||||
396 TTGAGAGGATGTGCTTGAAGTAATCAAGACTCTAATCCTGTAGACAAG 445
1892 CysLysLeuAspProAsnLysAlaTyrlleGlnIleThrTyrlValGluPr 1908

|||||
446 TGCAAATTAGATCCAAACAAGGCATATATTACAGATTACCTACGTGAGGCC 495
1908 oTyrrPheAspThrTyrgluMetLysAspArgIleThrTyrrPheAspLysA 1925
:::|||||
496 CTTCCTTGACACCTATGAGATGAAGACAGAGAACTACTTATT .GACAAA 544
1925 snTyrrAsnLeuArgArgPheMetTyrcysThrProPheThrLeuAspGly 1941
|||||
545 ATTATAATCTCCGGCGTTTCATGTACTGCACACCCTTCACTTAGATGGC 594
1942 ArgAlaHisGlyGluLeuHisGluGlnPheLysArgLysThrIleLeuTh 1958
|||||
595 CGTGCTCATGGGAGCTTCATGAACAATTCAACGGAAGACCATTTGAC 644
1958 rThrSerHisAlaPheProTyrlleLysThrArgValAsnValThrHis 1974
|||||
645 AACGTCATGTCTCTTCTTCTTACATTAAACAAGAGTCCATGTCACTCAC 694
1975 LysGluGluIleIleLeuThrProIleGluValAlaIleGluAspMetG1 1991
|||||
695 AAAGAAGAGATTATCTTAACACCAATTGAAGGTGCTATCCGAAGACTGCA 744
1991 nLysLysThrGlnGluLeuAlaPheAlaThrHisGlnAspProAlaAsp 2007
|||||
745 GAAACAGACCCGA.....ATGGGCTTGCACACACAATCTGTCCGGA 785
2008 ProLysMetLeu 2011
|||||
786 CCAAAGTAATG 797

seq_name: gb_htc:BC018075

seq_documentation_block:

LOCUS BC018075 1873 bp mRNA linear HTC 06-DEC-2001
DEFINITION Homo sapiens, clone IMAGE:4792506, mRNA.
ACCESSION BC018075
VERSION BC018075.1 GI:17390162
KEYWORDS HTC.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (03-DEC-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

REMARK
COMMENT NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabbs-r@mail.nih.gov

Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki
Toshiyuki and Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-shgc.stanford.edu>
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 32 Row: b Column: 21
This clone was selected for full length sequencing because it
passed the following selection criteria: Similarity but not
identity to protein
This clone has the following problem: frame shifted.

FEATURES
source location/Qualifiers
1..1873

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4792506"
/tissue_type="Brain, hypothalamus"
/clone_lib="NIH_MGC_96"
/lab_host="DH10B"
/note="Vector: pBluescript"
BASE COUNT 604 a 317 c 365 g 587 t
ORIGIN

alignment_scores:
Quality: 1199.00 Length: 229
Ratio: 5.259 Gaps: 1
Percent Similarity: 99.563 Percent Identity: 99.127

alignment_block:
US-09-737-246-2 x BC018075 ..

Align seg 1/1 to: BC018075 from: 1 to: 1873

1824 GlyTrp.GluArgMetPheGlyThrTyRpheArgValGlyPheTyrglyT 1840
||||| :|||
1170 GGATGGTAAGCGGATGTTGGCACCATATTTCGTGTGTTTATGGAA 1219
1840 hrlyspheglyaspheasppluIngluPheValTyrllysGluProAla 1856
|||||
1220 CCAAGTTCGGGATTTGGATGAACAAGAATTGTTTACAAGAGAGCCTGCA 1269
1857 IleThrLysLeuAlaGluIleSerHisArgLeuGluGlyPheTyrglyG1 1873
|||||
1270 ATAAACCAACTTGACAGATATCTCACAGATTGGAGGATTTTACGGAGA 1319
1873 uArgPheglyGluaspValAlaGluValIleLysaspSerAsnProVala 1890
|||||
1320 AAGATTGGAGAGAGATGGTGTGAAGTAATCAAGACTTAATCCTGTAG 1369
1890 splysCysLysLeuaspProAsnLysAlaTyrlleGlnIleThrTyrVal 1906
|||||
1370 ACAAGTGTAAATTAGATTCCTAACAAGGCATATATTACATATGATG 1419
1907 GluProTyRpheAspThrTyrgluMetLysAspArgIleThrTyRpheAs 1923
|||||
1420 GAGCCATACTTTGACACATATGAGATGAAGGACAGAAATCACCATTTCGA 1469
1923 pLysAsnTyRAsnLeuArgArgPheMetTyrcysThrProPheThrLeuA 1940
|||||
1470 CAAAATTACAAATCTTCGTCGATTTCATGTACTGTACACCCCTTACTTAG 1519
1940 spGlyArgAlaHisGlyLeuLeuHisGluGlnPheLysArgLysThrIle 1956
|||||
1520 ATGGCCGTGGCCATGGGGAACCTTCATGAACAATTCAAAAAGGAAGACCAT 1569
1957 LeuThrThrSerHisAlaPheProTyrlleLysThrArgValAsnValTh 1973
|||||
1570 CTGACTACGTCATCATGCCCTTCCCTTATATTAAACAAGGTCATGTAC 1619
1973 rHisLysGluGluIleIleLeuThrProIleGluValAlaIleGluaspM 1990
|||||
1620 TCATTAAGAAGAGATCATCTTAACACCAATGAGTTGCTATTGAGGACA 1669
1990 etGlnLysLysThrGlnGluLeuAlaPheAlaThrHisGlnAspProAla 2006
|||||
1670 TGCAGAAAAAGACACAGGAGTTGGCATTTCGACACATCAGGATCCCGCA 1719
2007 AspProLysMetLeuGlnMetValleuGlnGlySerValGlyThrThrVa 2023
|||||
1720 GACCCCAAAATGCTTCAGATGGTACTCCAGGAGATCTGTAGGCACACAGAT 1769
2023 lAsnGlnGlyProLeuGluValAlaGlnValPheLeuSerGluIleProS 2040
|||||
1770 GAATCAGGGGCGCTTTGGAAGTTGCCACAGGTTTCTGTCTGAATAACCTA 1819

2040 eraspProLysLeuPheArgHisHisAsnLysLeu 2051
|||||
1820 GTGACCCCAAGCTCTTCAGACATCATATAAACTG 1854

seq_name: gb_htc:BC018076

seq_documentation_block:
LOCUS BC018076 1873 bp mRNA linear HTC 06-DEC-2001
DEFINITION Homo sapiens, clone IMAGE:4792507, mRNA.
ACCESSION BC018076
VERSION BC018076.1 GI:17390164
KEYWORDS HTC.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (03-DEC-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

REMARK
COMMENT NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-shgc.stanford.edu>
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 32 Row: b Column: 20
This clone was selected for full length sequencing because it
passed the following selection criteria: Similarity but not
identity to protein
This clone has the following problem: frame shifted.

FEATURES
source location/Qualifiers
1..1873
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4792507"
/tissue_type="Brain, hypothalamus"
/clone_lib="NIH_MGC_96"
/lab_host="DH10B"
/note="Vector: pBluescript"

BASE COUNT 604 a 317 c 365 g 587 t
ORIGIN

alignment_scores:
Quality: 1199.00 Length: 229
Ratio: 5.259 Gaps: 1
Percent Similarity: 99.563 Percent Identity: 99.127

alignment_block:
US-09-737-246-2 x BC018076 ..

Align seg 1/1 to: BC018076 from: 1 to: 1873

1824 GlyTrp.GluArgMetPheGlyThrTyRpheArgValGlyPheTyrglyT 1840
||||| :|||
1170 GGATGGTAAGCGGATGTTGGCACCATATTTCGTGTGTTTATGGAA 1219
1840 hrlyspheglyaspheasppluIngluPheValTyrllysGluProAla 1856
|||||

```
1220 CCAAGTTCGGGGATTGGATGACAAGAATTTGTTTACAAGAGCCTGCA 1269
1857 ILeThrLysLeuAlaGluIleSerHisArgLeuGluGlyPheTyrGly 1873
1270 ATAACCAAACTTGACAGATATCTCACAGATTTGAGGGATTTTACGAGA 1319
1873 uArgPheGlyGluAspValValGluValIleLysAspSerAsnProVal 1890
1320 AAGATTGGAGAGAGATGTGGTTGAAGTAATCAAGACTCTAATCCTGTAG 1369
1890 sPlyScySLysLeuAspProAsnLysAlaTyrIleGlnIleThrTyrVal 1906
1370 ACAAGTGAATTAGATCCTTACACAAGGCATATATTTCAGATTACCTATGTG 1419
1907 GluProTyrPheAspThrTyrGluMetLysAspArgIleThrTyrPheAs 1923
1420 GAGCATACTTTGACACATATGAGATGAAGGACAGAAATCACCTATTTCGA 1469
1923 pLysAsnTyrAsnLeuArgPheMetTyrCysThrProPheThrLeuA 1940
1470 CAAAATTAACAATCTTCGTCGATTTCATGTACTGTACACCCTTACTTTAG 1519
1940 spGlyArgAlaHisGlyGluLeuHisGluGlnPheLysArgLysThrIle 1956
1520 ATGGCCGTGCCCATGGGGAAGCTTCATGAACAATTCAAAAGAAGACCATT 1569
1957 LeuThrThrSerHisAlaPheProTyrIleLysThrArgValAsnValTh 1973
1570 CTGACTACGTCCTCATGCGCTTTCCTTATATTAAACAAGGGTCAATGTCA 1619
1973 rHisLysGluGluIleIleLeuThrProIleGluValAlaIleGluAspM 1990
1620 TCATTAAGAAGAGATCATCTTAACACCAATTGAAGTTGCTATTGAGGACA 1669
1990 eTGlnLysLysThrGlnGluLeuAlaPheAlaThrHisGlnAspProAla 2006
1670 TGCAGAAAAAGACACAGAGAGTTGGCATTGGCAACACATCAGGATCCGCA 1719
2007 AspProLysMetLeuGlnMetValleuGlnGlySerValGlyThrThrVa 2023
1720 GACCCCAAAATGCTTCAGATGTTACTCCAGGATCTGTAGGCACACACAGT 1769
2023 lAsnGlnGlyProLeuGluValAlaGlnValPheLeuSerGluIlePro 2040
1770 GAATCAGGGGCGCTTGGAGTTGCCAGGTTTTCGTCTGAAATACCTA 1819
2040 eArgPProLysLeuPheArgHisHisAsnLysLeu 2051
1820 GTGACCCCAAGCTCTTCAGACATCATATAAACTG 1854
```

seq_name: gb_est2:BI689126

seq_documentation_block: 869 bp mRNA linear EST 18-SEP-2001
LOCUS BI689126
DEFINITION 603315930F1 NCI_CGAP_Mam6 Mus musculus cDNA clone IMAGE:5355798 5',
mRNA sequence.

ACCESSION BI689126
VERSION BI689126.1 GI:15651755
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Jeffrey Green M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LHAM1904 row: 1 column: 07
High quality sequence stop: 761.

FEATURES
source
1. 869
Location/Qualifiers

/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:5355798"
/clone_lib="NCI_CGAP_Mam6"
/sex="female, virgin"
/tissue_type="infiltrating ductal carcinoma"
/dev_stage="5 months"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies, Inc. Investigator
providing samples: Jeffrey Green, M.D., NIH"
BASE COUNT 223 a 213 c 237 g 196 t
ORIGIN

alignment_scores: Quality: 1193.00 Length: 271
Ratio: 4.553 Gaps: 4
Percent Similarity: 96.679 Percent Identity: 92.989

alignment_block:
us-09-737-246-2 x BI689126 ..
Align seg 1/1 to: BI689126 from: 1 to: 869

```
1284 MetIleSerGlnThrValAlaMetAlaIleAlaGlyThrSerValProG 1300
12 GTGATAAGCCAGACGCGTTGCCATGGCAATTGCAGAGAACATCGCTCCCTCA 161
1300 nLeuThrArgProGlySerPheLeuLeuThrSerThrSerGlyArgGlnH 1317
62 GGTGACACAGACCTGGCAGTTTCTTCCTCAGCTCAACGAGTGGCCGGCAGC 111
1317 iSThrThrPheSerAlaGluSerSerArgSerLeuLeuIleCysLeuLeu 1333
112 ACACACACCTTCTCAGCAGAGTCAAGCAGGACGCTTTTGATCTGCTGCTT 161
1334 TrpValLeuLysAsnAlaAspGluThrValleuGlnLysTrpPheThrAs 1350
162 TGGGTTCACAAGAATGCAGATGAACGTTCTGCAGAAATGGTTTACAGA 211
1350 pLeuSerValleuGlnLeuAsnArgLeuLeuAspLeuLeuTyrLeuCysV 1367
212 TCTCTCAGTCCTCCAGCTGAACCGGTTGTAGATCTGCTTACCTGTGTG 261
1367 aLserCysPheGluTyrLysGlyLysLys. ValPheGluArgMetAsnse 1383
262 TATCTTGCTTTTGAGTACAAAGGAAAAACGGTATTGTGAACGAATGAATAG 311
1383 rLeuThrPheLysLysSerLysAspMetArgAlaLysLeuGluGluAlaI 1400
312 TTTAACTTTTTAAGAAATCAAAAGACATGAGAGCCAAAGCTTTGAAGAGCCA 361
1400 lLeuGlySerIleGlyAlaArgGlnGluMetValArgArgSerArgGly 1416
362 TTCTGGGAGCATCGGTGCAGCAGGAAATGTTAGCGCGGAGCCGAGGG 411
1417 GlnLeuGluArgSerProSerGlySerAlaPheGlySerGlnGluAsnLe 1433
412 CAGCTCGAGAGAGAGCCCATCTGGAAGCGCTTTGGAGCCAGAAACCT 461
1433 uArgTrp.ArgLysAspMetThrHisTrpArgGlnAsnThrGluLysLeu 1449
462 GCGGTGAGAGCAAAAGACATGACTGCTGCGCTCAGAACTCAGAGAAACTC 511
```

1450 AsplysSerArgAlaGluIleGluHisGluAlaLeuIleAspGlyAsnIle 1466
|||||
512 GATTAAGTCAGAGCAGAGATGAGAACACGACACTGATGTGTAACCT 561
1466 uAlaThrGluAlaAsnLeuIleIleLeuAspThrLeuGluIleValValG 1483
|||||
562 GGCTACTGAAGCAAACTCATCATCTTAGACACCGCTAGACATCATCTCC 611
1483 InThrValSerValThrGluSerLysGluSerIleLeuGlyGlyValLeu 1499
|||||
612 AGACTGTGTCTGTAACGGAAATCCAAAGAGATATCTGGGTGTGCTA 661
1500 LysValLeuLeuHisSerMetAlaCysAsnGlnSerAlaValTyrLeuG1 1516
|||||
662 AAAGTGTCTGTACAGAGCATGGCTTGCAACCAAGT.GCAGTCTATCTGCA 710
1516 nHisCysPheAlaThrGlnArgAlaLeuValSerLysPheProGlu.Leu 1532
|||||
711 GCAGTGTCTTCGCCACGCAAGAGCCCTGGTCTCAAGTTCTGAGCTCTT 760
1533 LeuPheGluGluGluThrGluGluGlnCysAla.AspleuCysLeuArgLeu1 1549
:::
761 GTTCGAGGGGAAGCAGACCGACAGTGTGGCGGATCTGTGCTCCGCTTC 810
1549 euArg 1550
|||||
811 TCCGG 815
seq_name: gb_est2:BG872239
seq_documentation_block:
LOCUS BG872239 745 bp mRNA linear EST 29-MAY-2001
DEFINITION 602790848F1 NCI_CGAP_SG2 Mus musculus cDNA clone IMAGE:4922339 5',
mRNA sequence.
ACCESSION BG872239
VERSION BG872239.1 GI:14222779
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 745)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10841 row: O column: 12
High quality sequence stop: 737.
location/Qualifiers
1..745
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4922339"
/clone_1lb="NCI_CGAP_SG2"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: salivary gland; Vector: PCMV-SPORT6; Site_1:
NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo
dT. Average insert size 1.3 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT 219 a 167 c 196 g 163 t
ORIGIN

alignment_scores:

Quality:	1146.00	Length:	247
Ratio:	4.736	Gaps:	2
Percent Similarity:	97.976	Percent Identity:	95.142

alignment_block:

US-09-737-246-2 x BG872239 ..

Align seg 1/1 to: BG872239 from: 1 to: 745

1375 LysLysValPheGluArgMetAsnSerLeuThrPheLysLysSerLysAs 1391
|||||
3 AAAAAGGTAFTTGAACGAATGATAGTTTAACCTTTGAAGAAATCACAGA 52
1391 pMetArgAlaLysLeuGluGluAlaIleLeuGlySerIleGlyAlaArgg 1408
|||||
53 CATGAGAGCCAAAGCTTGAAGAGGCCATTCTGGGGAGCATCGGTGCCAGGC 102
1408 InGluMetValArgArgSerArgGlyGlnLeuGluArgSerProSergly 1424
|||||
103 AGGAATGTGTACGGCGGAGCCGAGGGGACGCTCGAGAGAGAGCCCATCTGGA 152
1425 SerAlaPheGlySerGlnGluAsnLeuArgTTrpArgLysAspMetThrHi 1441
|||||
153 AGCGCCTTTGGGAGCCCAAGAAAACCTGCGGTGGAGAAAGACATGACTCA 202
1441 sTrpArgGlnAsnThrGluLysLeuAspLysSerArgAlaGluIleGluH 1458
|||||
203 CTGGCGTCAAGACTCAGAGAAACTCGATAAGTCAAGAGAGAGATAGAAC 252
1458 isGluAlaLeuIleAspGlyAsnLeuAlaThrGluAlaAsnLeuIleIle 1474
|||||
253 ACGAAGCACTGATTTGATGGAAACCTGGCTACTGAAGCAAACTCATCATC 302
1475 LeuAspThrLeuGluIleValValGlnThrValSerValThrGluSerly 1491
|||||
303 TTAGACACGCTAGAGATCATCTGTCAGACTGTCTCTGTACGGAATCCAA 352
1491 sGluSerIleLeuGlyGlyValLeuLysValLeuLeuHisSerMetAlaC 1508
|||||
353 AGAGATATCTGGGTGTGTGTTAAAGTGTCTGTACAGACATGCGCT 402
1508 ysAsnGlnSerAlaValTyrLeuGlnHisCysPheAlaThrGlnArgAla 1524
|||||
403 GCAACCAAAAGTGCAGTCTATCTGCACACTGCTTGGCCACGACAGAGCC 452
1525 LeuValSerLysPheProGluLeuLeuPheGluGluThrGluGlnCyl 1541
|||||
453 CTGGTCTCAAGTTTCTCTGAGCTCTTGTGAGGAGAGACAGACAGCAGTG 502
1541 sAlaAspLeuCysLeuArgLeuLeuArgHisCysSerSerSerIleGlyT 1558
|||||
503 TGGGATCTGTGCTCCGGCTTCTCCGTCAATTGCAGTAGCAGCATCAGTA 552
1558 hrIleArgSerHisProSerAlaSerLeuTyrLeuLeuMetArgGlnAsn 1574
|||||
553 CAATCCGGTGGCAGCGCTAGTGCCTCCCTTTAACCCTCATGAGCGACGAAC 602
1575 PheGluIleGlyAsnAsnPheAlaArgValLysMetGlnValPrometSe 1591
|||||
603 TTTGAGATCGGGAATAACTTTTGCCAGAGTGAATAATGCAAGTGACGATGTC 652
1591 rLeuSerSerLeuValGlyThrSerGlnAsnPheAsn.GluGluPheLeu 1607
|||||
653 ACTGTCTCTTAGTGGGTAGCTAGTCTCAGAAATTTTAATCGAAGAAATCTTC 702
1608 Arg.ArgSerLeuLysThrIleLeuThrTyrAlaGlu 1619
|||||
703 AGACCGCTCTTAATAAACTATTTTGACATATGCTGAA 739
seq_name: gb_hlc:AK013336
seq_documentation_block:

LOCUS	AK013336	636 bp	mRNA	linear	HTC 19-JAN-2002
DEFINITION	Mus musculus 10, 11 days embryo whole body cDNA, RIKEN full-length enriched library, clone:2810453G12:homolog to BA165F24.1.1 (NOVEL PROTEIN SIMILAR TO DROSOPHILA CG6630 AND CG11376, KIAA1058, RAT TRG (ISOFORM 1)) (FRAGMENT), full insert sequence.				
ACCESSION	AK013336				
VERSION	AK013336.1	GI:12850636			
KEYWORDS	HTC; CAP trapper. Mus musculus (strain:C57BL/6J) 10, 11 days embryo cDNA to mRNA, clone_lib:RIKEN full-length enriched mouse cDNA library clone:2810453G12.				
SOURCE					
ORGANISM	Mus musculus				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
AUTHORS	Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.				
TITLE	Normalization and subtraction of cap-trapper-selected CDNAs to prepare full-length cDNA libraries for rapid discovery of new genes				
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)				
MEDLINE	20499374				
PUBMED	11042159				
REFERENCE	3 (sites)				
AUTHORS	Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Wataniki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.				
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer				
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)				
MEDLINE	20530913				
PUBMED	11076861				
REFERENCE	4 (sites)				
AUTHORS	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.				
TITLE	Functional annotation of a full-length mouse cDNA collection				
JOURNAL	Nature 409, 685-690 (2001)				
REFERENCE	5 (bases 1 to 636)				
AUTHORS	Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A., Arakawa,T., Baldarelli,R., Bono,H., Brownstein,M., Bult,C., Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hill,D., Hiramoto,K., Hiraoka,T., Horii,F., Hume,D., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Quackenbush,J., Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Schriml,L., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yamanaka,I., Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.				
TITLE	Direct Submission				
JOURNAL	Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Gemomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsr.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)				
COMMENT	Please visit our web site (http://genome.gsc.riken.go.jp/) for further details. cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken				

Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues. First strand cDNA was primed with a primer
[5' GAGAGAGAGAAGATCCAAAGACTCTTTTCTTTTTTTTTTTVN 3'], cDNA was
prepared by using trehalose thermo-activated reverse transcriptase
and subsequently enriched for full-length by cap-trapper. cDNA went
through one round of normalization to Rot = 7.5 and subtraction to
Rot = 37.5. Second strand cDNA was prepared with the primer adapter
of sequence [5'
GAGAGAGACATTCGCAGTTAATTAAATTATCCCCCCCCCCC 3']. cDNA was cleaved
with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI.
Host: SOLR.

```

FEATURES             location/qualifiers
source               . 1..636
                    /organism="Mus musculus"
                    /strain="C57BL/6J"
                    /db_xref="MGD:MGI:1896206"
                    /db_xref="taxon:10090"
                    /clone="2810453G12"
                    /clone_lib="RIKEN full-length enriched mouse cDNA library"
                    /dev_stage="10, 11 days embryo"
                    /dev_...>636
                    /note="data source:SPRR, source key:Q9H1Q3, evidence:ISS
homolog to BA165F24.1.1 (NOVEL PROTEIN SIMILAR TO
DROSOPHILA CG6630 AND CG11376, KIAA1058, RAT TRG (ISOFORM
1)) (FRAGMENT)
putative"
/codon_start=1
/protein_id="BAB28798.1"
/db_xref="GI:12850637"
/translation="MKDRITYFDKNYNLRFMYCTPTFLDGRANGELHEQFKRKILT
TSHAFPIKTRVNVTHKEEIIITPEVAIEDMOKKTOLAEFATHODPADPKMLQMVLO
GSVGTVNQGPLEVAQVFLEISIRPGDPKLFRRHNKLRLCFKEDFTKRCEDALR"

BASE COUNT          202 a      141 c      133 g      160 t
ORIGIN

alignment_scores:
    Quality: 1106.00           Length: 211
    Ratio: 5.242              Gaps: 0
Percent Similarity: 100.000   Percent Identity: 98.578

alignment_block:
US-09-737-246-2 x AK013336 ..

Align seg 1/1  to: AK013336 from: 1 to: 636

1856 Alailethrllysleualaagluileserhisargleuglyphetyrgl 1872
|||||
2 GCATAACCAAACTCGCACAGATTCTCACAGATTGGAGGATTTATGG 51

1872 Ygluargpheglygluaspyalvalglualillelysaspserasnprov 1889
|||||
52 AGAAAGATTGGAGAGGATGTGCTTGAAGTAATCAAGACTCTAATCTTG 101

1889 alaspyscylsylsleuaspproasnlysalaatyrtleglnilethtyr 1905
|||||
102 TAGACAAGTGCAAAATTAGATCCAACAAGCATATATTCAGATTACCCTAC 151

1906 valgluprotyrpheasprthrtyrglumetlysaspargillethtyrph 1922
|||||
152 GTGAGGCCCTTCTTTGACACCTATGAGATGAAGGACAGAATCACTTATTT 201

1922 easplysasntyrasnleuatgargphemetyrcysthrprophetnrl 1939
|||||
202 TGACAAAAAATTATATCTCCGGCGTTTCATGTACTGCACACCCCTTCACTT 251

1939 euaspglyargalahisglygluleuhisgluginphelysarqlysthr 1955
|||||
252 TAGATGCCCGTCTCATAGGGAGCTTCATGAACAATTCAAACGGAGACC 301

1956 lleuthrthrsernhisalarpherotyrtlleystthrarqvalasnva 1972
|||||
```



```
302 ATCTGACACGTCATGCCTTCTTACATTAACAAGAGTCATGT 351
1972 lThrHisLysGluIleIleIleLeuThrProIleGluValAlaIleGluA 1989
      |||||||
352 CACTCACAAAGAAGAGATTAATCTTAACACCATTGAAGTTGCTATCGAAG 401
1989 sPmetGlnLysLysThrGlnGluLeuAlaPheAlaThrHisGlnAspPro 2005
      |||||||
402 ACATGCAGAAAAAGACTCAGGAATTGGCATTTGCCAACACATCAGGATCCA 451
2006 AlaAspProLysMetLeuGlnMetValLeuGlnGlySerValGlyThrTh 2022
      |||||||
452 GCAGACCCCAAAATGCTCAAAATGCTCCAGGGGTCTGTAGAACACAC 501
2022 rValAsnGlnGlyProLeuGluValAlaGlnValPheLeuSerGluIleP 2039
      |||||||
502 TGTAAACCCAGGACCTCTGGAGAGTTGCCACAGTCTTCTGTGAATAC 551
2039 rOSerAspProLysLeuPheArghHisHisAsnLysLeuArgLeuCysPhe 2055
      ||:::|||||
552 CTGGTGACCCGGAAGCTTTCAGACATCACACACAGCTGCGCTTGTGTTT 601
2056 LysAspPheThrLysArgCysGluAspAlaLeu 2066
      |||||||
602 AAGGACCTTACTAAAGGTGTGAGGATGCCTTA 634
```

seq_name: gb_est2:BI770146

seq_documentation_block:

LOCUS BI770146 796 bp mRNA linear EST 25-SEP-2001
DEFINITION 603053332F1 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5203063 5',
mRNA sequence.

ACCESSION BI770146 GI:15761724
VERSION BI770146.1
KEYWORDS EST.

SOURCE human.
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 796)
NIH-MGC http://mgc.nci.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov

Tissue Procurement: Life Technologies, Inc.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

plate: LLAM1508 row: p column: 08

High quality sequence stop: 793.

FEATURES

source

1..796

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:5203063"

/clone_11b="NIH_MGC_122"

/lab_host="DH10B"

/note="Organ: pooled lung and spleen; Vector: pCMV-SPORT6;
Site_1: NotI; Site_2: EcoRV (destroyed); RNA source

anonymous pool of 24 week female lung, 16 week female

spleen, and 20-22 week male spleens. Library is oligo-dT

primed and directionally cloned (EcoRV site is destroyed

upon cloning). Average insert size 1.4 kb, insert size

range 1-3 kb. Library is normalized and enriched for

full-length clones and was constructed by C. Gruber

(Invitrogen). Research Genetics tracking code 026. Note:

this is a NIH_MGC Library."

BASE COUNT 188 a 228 c 231 g 148 t 1 others

ORIGIN

alignment_scores: Quality: 1104.00 Length: 250
Ratio: 4.738 Gaps: 1
Percent Similarity: 93.200 Percent Identity: 82.800

alignment_block:
US-09-737-246-2 x BI770146 ..

Align seg 1/1 to: BI770146 from: 1 to: 796

```
1836 GlyPheTyrgLyThrLysPheGlyAspLeuAspGluGlnGluPheValTy 1852
      |||||||
1 GGTCTCTACGGCGCCGACCTTCGGTGACCTGGATGAGCAGGAGTTGTGTA 50
1852 rLysGluProAlaIleThrLysLeuAlaGluIleSerHisArgLeuGluG 1869
      |||||||
51 CAAGGAGCCATCGATCAGCAGAGCTGGCAGAGATCTCACACCGGCTGGAGG 100
1869 lYPheTyrgLyGluArgPheGlyGluAspValValGluValIleLysAsp 1885
      |||||
101 AGTTCTACACGGAGAGATTGGCCGACGACGCTCGTTGAGATTATCAAGAC 150
1886 SerAsnProValAspLysCysLysLeuAspProAsnLysAlaTyrlleGl 1902
      |||||||
151 TCTAACCTGTGGACAAAGTCCAAAGCTTGACTCACAACAAAGCCCTACATCCA 200
1902 nIleThrTyrrValGluProTyrrPheAspThrTyrgLumetLysAspArgI 1919
      |||||||
201 GATCACGTAATGTGGAACCGTACTTGTGATACCTACGAGCTCAAGACCGGG 250
1919 lEThrTyrrPheAspLysAsnTyrrAsnLeuArgArgPheMetTyrcystThr 1935
      :::|||||
251 TGACCTACTTGTGACCGCAACTATGGGCTTCGCACATTCCTGTTGACAG 300
1936 ProPheThrLeuAspGlyArgAlaHisGlyGluLeuHisGluGlnPheLy 1952
      |||||||
301 CCGTTCACGCGCGGATGGGCGCGCACAGGGGAGAGCTGCCCGACACACAA 350
1952 sArgLysThrIleLeuThrThrSerHisAlaPheProTyrlleLysThra 1969
      |||||||
351 GCGTAAGACGCTGCTCAGACCGACCGCCTTCCCTACATCAAGACTC 400
1969 rGValAsnValThrHisLysGluGluIleIleLeuThrProIleGluVal 1985
      ||:::||||
401 GCATCCGTGTGTGCCACCGGAGGAGACGGTGTGACGCGCACATGGAGGTG 450
1986 AlaIleGluAspMetGlnLysLysThrGlnGluLeuAlaPheAlaThrHi 2002
      |||||||
451 GCCATCGAGGACATGCAGAAAGACACGGGAGCTGGCCTTGCCACCGGA 500
2002 sGlnAspProAlaAspProLysMetLeuGlnMetValLeuGlnGlyServ 2019
      :|||||
501 GCAGGACCCACGAGATGCTAAGATGCTACAGATGCTGCTTCAGGGCTCTG 550
2019 aGlyThrThrValAsnGlnGlyProLeuGluValAlaGlnValPheLeu 2035
      |||||
551 TAGGGCCACCGCTGAACACCAAGGTCCTCCCTGAGAGTGGCCCAAGGTGTTT 600
2036 SerGluIleProSerAspProLysLeuPheArghHisHisAsnLysLeuAr 2052
      :::|||||
601 GCAGAGATCCCGGAAGACCCCAAGCTTTCGGGCATCACACAAATTGGCG 650
2052 gLeuCysPheLysAspPheThrLysArgCysGluAspAlaLeuArgLys. 2068
      |||||||
651 GCTCTGCTTCAAGACTTCTGCAAGAAATGTGAGGATGCGGTGCGGAAC 700
2069 AsnLysSerLeuIleGlyProValGlnLysGluTyrgLInArgGluLeu 2084
      |||||||
701 AATAAGGCCCTGATTTGGGCGGACACAGAGAGATACACCGGTGAGCTG 748
```

seq_name: gb_est2:BF309712

```
seq_documentation_block:
LOCUS      BF309712                898 bp    mRNA       linear   EST 21-NOV-2000
DEFINITION 601891854F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:4137732 5',
            mRNA sequence.
ACCESSION  BF309712
VERSION    BF309712.1  GI:11257158
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 898)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Ling Hong/Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LNL at: image.lnl.gov
            Plate: L1CM1045 row: k column: 13
            High quality sequence stop: 732.
FEATURES
    source
        1..898
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="IMAGE:4137732"
            /clone_1lb="NIH_MGC_17"
            /tissue_type="rhabdomyosarcoma"
            /lab_host="DH10B (phage-resistant)"
            /note="Organ: muscle; Vector: pOTB7; Site_1: EcoRI;
            Site_2: XhoI; cDNA made by oligo-dT priming.
            Directionally cloned into EcoRI/XhoI sites using the
            following 5' adaptor: GGCACGAG(G). Size-selected >500bp
            for average insert size 1.8kb. Library constructed by
            Ling Hong in the laboratory of Gerald M. Rubin (University
            of California, Berkeley) using ZAP-cDNA synthesis kit
            (Stratagene) and Superscript II RT (Life Technologies)."
```

```
BASE COUNT      276 a      182 c      228 g      212 t
ORIGIN
alignment_scores:
    Quality: 1092.00          Length:      282
    Ratio: 4.233              Gaps:      4
    Percent Similarity: 91.489    Percent Identity: 87.234
alignment_block:
US-09-737-246-2 x BF309712  ..
Align seg 1/1  to: BF309712  from: 1  to: 898
1378  phegluArgmetAsnSerLeuThrphelyLysSerLysAspmetArgAl 1394
      |||||
3    TTTGAACGAATGAATAGCTTGACCTTTAAGAAATCAAAAAGACATGAGAGC 52
1394  aLysLeuGlUGlUAAlaIleLeuGlYSerIleGlyAlaAArgGlNGluMetV 1411
      |||||
53  AAAGCTTGAAGAGACTAATTCTTGGGAGCATAGTGCCAGGCAAGAAATGG 102
1411  aLArgARgSerARgGlYGlNleuGlUAArgSerProSerGlySerAlaPhe 1427
      |||||
103  TACGGCGAAGCCGAGGACAGCTCGAGAGAACCCATCTGGAAGTGCCCTTT 152
1428  GlYSerGlNGluAsnLeuARgTTPARgLYsAspMetThrHisTrrArgGl 1444
      |||||
153  GGAAGTCAAGAAAAATTGAGCTGGAGGAAGATATGACTCACTGGCGTCA 202
1444  nAsnThrgLUlYsLeuAspLYsSerARgAlaGlUleGlUHisGlUAAlaL 1461
```

```
|||||
203  AAACACAGACAGAGCTTGACAAATCAAGAGCAGAGATTGAACACGAAGCAC 252
1461  euILeAspGlYAsnLeuAlaThrGlUAAlaAsnLeuIleIleLeuAspThr 1477
      |||||
253  TGATTGATGGAAACCTGGCTACAGAACCAAACTAATCATTTTAGATACA 302
1478  LeuGlUleValValGlNThrValSerValThrgLUserLYsGlUserI 1494
      |||||
303  TTAGAGATTGT.GTTCAGACCGTTTCTGTAAACGGAATCCAAAGAGAGCAT 351
1494  eLeuGlYGlYValLeuLysValLeuLeuHisSerMetAlaCysAsnGlns 1511
      |||||
352  TCTTGGTGAGTGCTTAAGTGCCTACTACACAGCATGGCCCTGTAACCAA 401
1511  eAlaValTYrLeuGlNHisCysPheAlaThrglNArgAlaLeuValSer 1527
      |||||
402  GTGCAGTTATCTACACACAGTGTTCCTACACAGAGAGCCCTTGCTTCA 451
1528  LysPheProGlUleuLeuPheglUGlUGlUThrglUGlNcysAlaAspLe 1544
      |||||
452  AAGTTTCCTGAACCTCTTATTTGAAGAAGACAGACAGAGCAGTGTGCTGATT 501
1544  uCysLeuARgLeuLeuARgNHisCysSerSerSerIleGlyThrIleArgS 1561
      |||||
502  ATGCCCTCAGGCTTCTCCGACACAGTAGCAGTAGCATCGGTACAATACGGT 551
1561  eRHisProSerAlaSerLeuTYrLeuLeuMetARgAlnAsnPheglUle 1577
      |||||
552  CACACGCCAGTGCTCCCTTACCTACTAATGAGGCAAAACTT.GAGATC 600
1578  GLYAsnAsnPheAlaARgValLYsMetGlNValPrometSerLeuSerSe 1594
      |||||
601  GGGAATTACTTTGCCAGGGTTAAATGCAGGTAACAATGTCATATCCTC 649
1594  rLeuValGlYThrSerGlNAsnPheAsnGlUGlUPheLeuARgSerL 1611
      |||||
650  .CTGGTGGGCACATCTCAGGAATTTTAATGAAGAATCTTAAGACGTTCTC 699
1611  euLYsThrIleLeuThr.TYrAlaGlUGlUAspLeuGlUleuARgGlUth 1627
      ||:::
700  TAAGACTATTAATTGACACATAATGCTGAGAAAGATCTGGCAATGGCGGGAG 749
1627  rThrPheProAspGlNValGlN.AsPLeuValPheAsnLeuHisMetIle 1643
      ::|||
750  CACCTTGCCGGATCAAGGCCAGAGATCGGGTT.....CACACCAC 790
1644  LeuSerAspThrValLYsMetLYsGlUHisGlNGlUAspPro 1657
      ::|||
791  ATGAGTTTCTGTGTCACGTGTAGA.....GAGGACCCC 823
```

1
2
3
4

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 16, 2002, 08:33:13 ; Search time 646.39 Seconds
(without alignments)
18136.252 Million cell updates/sec

Title: US-09-737-246-1

Perfect score: 6828

Sequence: 1 gtcgcgcgcgcgcgcagcgc.....ctcgtgcgcgcgcgcgcctc 6828

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_032802.*
1: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	6828	100.0	6828	22	AAH41934	Human CLASP-3 cDNA
2	4123	60.4	4143	22	AAH41911	Preliminary human
3	4123	60.4	4144	22	AAH41912	Human CLASP-3 cDNA
4	4123	60.4	4200	22	AAH41953	CLASP-3 nucleotide
5	3044.8	44.6	3173	22	AAH41952	CLASP-3 nucleotide
6	2132.8	31.2	6372	22	AAH43851	Human CLASP-7 enco
7	1749	25.6	7215	22	AAS07373	Human cDNA encodin
8	1079	15.8	4027	22	AAS07382	Human DNA associat
9	919.6	13.5	2148	22	AAH43850	Preliminary human

C	10	841.6	12.3	2014	22	AAS07381	Human DNA associat
	11	621.6	9.1	9389	23	ABL17736	Drosophila melanog
	12	565	8.3	1652	22	AAH99572	Human protein enco
	13	459.8	6.7	2427	22	AAI61331	Human polynucleoti
	14	434.4	6.4	2433	22	AAI59545	Human polynucleoti
	15	370	5.4	1834	24	AAS62625	CDNA sequence #412
	16	334.4	4.9	933	22	AAH41933	Human CLASP-3 geno
	17	324.6	4.8	1566	22	AAH17340	Human cDNA sequenc
	18	321.2	4.7	5589	23	ABL17737	Drosophila melanog
	19	304.8	4.5	2610	22	AAF32647	Human cDNA encodin
	20	304.8	4.5	6454	22	AAS08334	Human cDNA encodin
	21	301.6	4.4	3472	22	AAS08358	Human cDNA encodin
	22	298.4	4.4	4393	22	AAS08357	Human cDNA encodin
C	23	281	4.1	365	14	AAQ60222	Human brain Expres
	24	245.8	3.6	1273	21	AAC93501	Human secreted pro
	25	244.4	3.6	4108	22	ABA18725	Human nervous syst
	26	216.2	3.2	260	22	AAH41940	Human CLASP-3 geno
	27	214.6	3.1	3899	22	AAF32675	Human cDNA encodin
	28	207.8	3.0	5688	21	AAC87969	Human CLASP relate
	29	207.8	3.0	5688	21	AAA14825	DNA encoding a hum
	30	207.8	3.0	7277	24	AA518951	Human cDNA encodin
	31	196.4	2.9	1605	22	AA507383	Human DNA associat
	32	190.4	2.8	996	23	AA587322	DNA encoding novel
	33	190.4	2.8	2036	22	AAH14086	Human cDNA sequenc
	34	183.4	2.7	5214	21	AAC87968	Mouse CLASP-1 nucl
	35	183.4	2.7	5214	21	AAA14824	DNA encoding a mur
	36	180.2	2.6	441	22	AA184302	Human polynucleoti
	37	178.2	2.6	220	22	AAH41937	Human CLASP-3 geno
	38	174.6	2.6	586	21	AAC75563	Human ORFX ORF1118
	39	171.6	2.5	2165	21	AAC76949	Human ORFX ORF2504
	40	168	2.5	211	22	AAH41945	Human CLASP-3 geno
	41	167.4	2.5	221	22	AAH41949	Human CLASP-3 geno
	42	154.6	2.3	201	22	AAH41943	Human CLASP-3 geno
	43	153.6	2.2	6816	21	AAC74524	Human ORFX ORF79 p
	44	153.6	2.2	7506	22	AAJ19118	Angiogenesis assoc
	45	148	2.2	981	23	AA570846	DNA encoding novel

ALIGNMENTS

RESULT	1	
AAH41934	standard; cDNA; 6828 BP.	
XX	AAH41934;	
AC	AAH41934;	
XX		
DT	30-AUG-2001 (first entry)	
XX		
DE	Human CLASP-3 cDNA sequence SEQ ID NO:1.	
XX		
KW	Human; cadherin-like asymmetry protein; CLASP; immune response;	
KW	cell surface molecule; transmembrane protein; immunosuppressive; vaccine;	
KW	antiinflammatory; antiarthritic; antianaemic; dermatological; uropathic;	
KW	ophthalmological; antirheumatic; nephrotropic; antithyroid; antidiabetic;	
KW	neuroprotective; antiasthmatic; antibacterial; antisense therapy;	
KW	gene therapy; chromosome 1p31.1; ss.	
XX		
OS	Homo sapiens.	
XX		
PN	WO200142297-A2.	
XX		
PD	14-JUN-2001.	
XX		
PF	13-DEC-2000; 2000WO-US34171.	
XX		
PR	13-DEC-1999; 99US-0170453.	
PR	14-JAN-2000; 2000US-0176195.	
PR	14-FEB-2000; 2000US-0182296.	
PR	11-APR-2000; 2000US-0196267.	
PR	11-APR-2000; 2000US-0196460.	
PR	11-APR-2000; 2000US-0196527.	
PR	11-APR-2000; 2000US-0196528.	

PR 11-APR-2000; 2000US-0547276.
PR 13-OCT-2000; 2000US-0240503.
PR 13-OCT-2000; 2000US-0240508.
XX
PA (ARBO-) ARBOR VITA CORP.
XX
XX
PI Lu P, Garman JD, Candia AF;
XX
DR WPI: 2001-375003/39.
DR P-PSDB; AAB99495.
XX
PT Novel isolated cadherin-like asymmetry protein (CLASP) -3, useful for
PT useful in preventing or treating a CLASP-3-mediated disease e.g.
PT autoimmune disease
XX
XX
PS Claim 1; Fig 6; 189pp; English.

XX
CC The present invention describes an isolated polypeptide (I) comprising a
CC nucleotide (nt) sequence (SIa) that has at least 90% identity to the
CC sequence given in AAH41934 and is immunologically cross-reactive with
CC the derived amino acid (aa) sequence (SIb) given in AAB99495 or shares a
CC biological function with native CLASP-3. (I) has immunosuppressive,
CC antiinflammatory, antiarthritic, antianemic, dermatological, uropathic,
CC ophthalmological, antirheumatic, nephrotropic, antithyroid, antidiabetic,
CC neuroprotective, antiasthmatic and antibacterial activities, and can be
CC used in antisense therapy, vaccine production and gene therapy. CLASP-3
CC related sequences can be used in preventing or treating a CLASP-3
CC mediated disease, preferably an autoimmune disease by inhibiting an
CC immune response. The autoimmune disease is caused or exacerbated by
CC increased activity of T helper 1 lymphocytes (TH1). AAH41893 to AAH41953
CC and AAB99491 to AAB99507 represent sequences which are used in the
CC exemplification of the present invention. CLASP-3 is localised in the
CC chromosome location 1p31.1.
XX
SQ Sequence 6828 BP; 2082 A; 1384 C; 1432 G; 1930 T; 0 other;

Query Match 100.0%; Score 6828; DB 22; Length 6828;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 6828; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gtcgcgtgcgcgcgcagcagccatgcccgcgcgcgcgccttgcgccagaagatcagcaga 60
Db 1 gtcgcgcgtgcgcgcgcagcagccatgcccgcgcgcgcgccttgcgccagaagatcagcaga 60
QY 61 acggtgycagccggaagttaggaagcagatctccgcgaatatagtgttctccccaactg 120
Db 61 acggtgycagccggaagttaggaagcagatctccgcgaatatagtgttctccccaactg 120
QY 121 ctcaaaaaccttaatatgtgtgcaatatatacccatcacacacagtgccctaccgaa 180
Db 121 ctcaaaaaccttaatatgtgtgcaatatatacccatcacacacagtgccctaccgaa 180
QY 181 gcaagtagatccagtgattgtggaagattaccttactcatccttggctgtgtattct 240
Db 181 gcaagtagatccagtgattgtggaagattaccttactcatccttggctgtgtattct 240
QY 241 gggccttaagggtattgtattgaatttcctccagatgatattgaagttgtttatagtcct 300
Db 241 gggccttaagggtattgtattgaatttcctccagatgatattgaagttgtttatagtcct 300
QY 301 cgggactgcagaactctgtttcagctgtacctggaagaagtgaatgataccacatgtt 360
Db 301 cgggactgcagaactctgtttcagctgtacctggaagaagtgaatgataccacatgtt 360
QY 361 agagactgtataagaagtatacagaagactgggcaattgtcatcagaaaatatcataaa 420
Db 361 agagactgtataagaagtatacagaagactgggcaattgtcatcagaaaatatcataaa 420
QY 421 ttgggaacagatattaatcccaatacattagataaaacagaaagaaagccttggttg 480
Db 421 ttgggaacagatattaatcccaatacattagataaaacagaaagaaagccttggttg 480

QY 481 ccaaaacaagtttttgaatctgtatgaagctccaagatgycacaagctaccagatgatcaa 540
Db 481 ccaaaacaagtttttgaatctgtatgaagctccaagatgycacaagctaccagatgatcaa 540
QY 541 gatgaccttaaaagaagttcaatgtcacaatagatgatacccaagggttagctggcctgt 600
Db 541 gatgaccttaaaagaagttcaatgtcacaatagatgatacccaagggttagctggcctgt 600
QY 601 agtactttagacttgaaaaaattcaacttctgtatgcttgcctcccaattacttgatcga 660
Db 601 agtactttagacttgaaaaaattcaacttctgtatgcttgcctcccaattacttgatcga 660
QY 661 actccaaatgaagaatagaccgttcagaatgatgaccaaaagaaatccaaccgtcacaaa 720
Db 661 actccaaatgaagaatagaccgttcagaatgatgaccaaaagaaatccaaccgtcacaaa 720
QY 721 gaacttttgccttgcatccatccaccagatgaggaagaaccaatagaaacggcttagtgtt 780
Db 721 gaacttttgccttgcatccatccaccagatgaggaagaaccaatagaaacggcttagtgtt 780
QY 781 cctgatatacccaagaacatttgcgtcaaaagacttctgttaaaatgcttatcactcaag 840
Db 781 cctgatatacccaagaacatttgcgtcaaaagacttctgttaaaatgcttatcactcaag 840
QY 841 ttggaattggaattgaaccatcttgcgaagttggcttataatgatagtcagaagaaag 900
Db 841 ttggaattggaattgaaccatcttgcgaagttggcttataatgatagtcagaagaaag 900
QY 901 aaaaagattcagaaaaactttatattgcacctaatctgagcagatgaaaggttgtta 960
Db 901 aaaaagattcagaaaaactttatattgcacctaatctgagcagatgaaaggttgtta 960
QY 961 cgtccacatgtaccacactgctgccattactaccctgccaagatcagcaattttctatc 1020
Db 961 cgtccacatgtaccacactgctgccattactaccctgccaagatcagcaattttctatc 1020
QY 1021 acttatccttcccaagatgttttctgttaataaagctagaaaaaagctctacagcaagga 1080
Db 1021 acttatccttcccaagatgttttctgttaataaagctagaaaaaagctctacagcaagga 1080
QY 1081 gacattggagagtgtygcagaaacatatatgtatttcaaaagaagcagatgccaaccaagaat 1140
Db 1081 gacattggagagtgtygcagaaacatatatgtatttcaaaagaagcagatgccaaccaagaat 1140
QY 1141 aaagaaaaactggaagaactgaagagtcgaagcagatcagtttgcctcaaaagacttgggaaa 1200
Db 1141 aaagaaaaactggaagaactgaagagtcgaagcagatcagtttgcctcaaaagacttgggaaa 1200
QY 1201 tatcgcatgccttttgccttgcgtggaactgccatttaatgaatatgttagcagtgctgg 1260
Db 1201 tatcgcatgccttttgccttgcgtggaactgccatttaatgaatatgttagcagtgctgg 1260
QY 1261 agtttgaaagagattctacagaagtagaataacagtaactgagagaacgaaaggtcttgg 1320
Db 1261 agtttgaaagagattctacagaagtagaataacagtaactgagagaacgaaaggtcttgg 1320
QY 1321 tcagagagaggaattctagtatgtttgycagagcagatcacacttgaaggacacaagaagtga 1380
Db 1321 tcagagagaggaattctagtatgtttgycagagcagatcacacttgaaggacacaagaagtga 1380
QY 1381 gatgatgcttctgaacttgcagagcttctgcaccagctactctcacagtgacaatttttct 1440
Db 1381 gatgatgcttctgaacttgcagagcttctgcaccagctactctcacagtgacaatttttct 1440
QY 1441 aagcaggaagaggaacccgttaagtgaatgaagatctctacaactccttgcgtatagaga 1500
Db 1441 aagcaggaagaggaacccgttaagtgaatgaagatctctacaactccttgcgtatagaga 1500
QY 1501 aggcacatcttctgtcttaacggcgaactaagacctattacagctcagctcaagatagacatt 1560
Db 1501 aggcacatcttctgtcttaacggcgaactaagacctattacagctcagctcaagatagacatt 1560
QY 1561 tctcccgacactgaaaaatcccatattattgcttaactccggagctgcttcaagtgaagctt 1620

Db 1561 tctccgcacctgaaaaatccccattattgtccctaactccgagctgcttcaagtgaaagctt 1620
QY 1621 tacccctgacagtagagttagacctaccagagaaatcttagagttcccgcaaggatgtt 1680
Db 1621 tacccctgacagtagagttagacctaccagagaaatcttagagttcccgcaaggatgtt 1680
QY 1681 tatgttccaacactacttacagaaatcttctctacatatataccctcaagctttaattt 1740
Db 1681 tatgttccaacactacttacagaaatcttctctacatatataccctcaagctttaattt 1740
QY 1741 gcccaatcgtaaggttctgtctagaataatacaagtgaaagtcgaagttatgttagagag 1800
Db 1741 gcccaatcgtaaggttctgtctagaataatacaagtgaaagtcgaagttatgttagagag 1800
QY 1801 gatcccaagcaatgccaatgccggtlaatcttggtaaatctagctgttcagaatttccaag 1860
Db 1801 gatcccaagcaatgccaatgccggtlaatcttggtaaatctagctgttcagaatttccaag 1860
QY 1861 gaagcctatacagccgtagtatatacatacaaggtcctcgatatttcatagaagaatcaag 1920
Db 1861 gaagcctatacagccgtagtatatacatacaaggtcctcgatatttcatagaagaatcaag 1920
QY 1921 gttaaagcttcctgtactttaactgacacatcaactgtgcttttacttttatcatgtt 1980
Db 1921 gttaaagcttcctgtactttaactgacacatcaactgtgcttttacttttatcatgtt 1980
QY 1981 agttgtcaacaaaaaataatactactcctcttgaacaacccaagttgatacataatgatacca 2040
Db 1981 agttgtcaacaaaaaataatactactcctcttgaacaacccaagttgatacataatgatacca 2040
QY 2041 atgcttcagaatgagcggttgaagactgagccagtttctgtccagctcatctygaaaaa 2100
Db 2041 atgcttcagaatgagcggttgaagactgagccagtttctgtccagctcatctygaaaaa 2100
QY 2101 ccaccacaggttatctgtactgtctcctgaggttcctctacccctgagcatgaaatgggta 2160
Db 2101 ccaccacaggttatctgtactgtctcctgaggttcctctacccctgagcatgaaatgggta 2160
QY 2161 gataatcacaaagtggttttaatggtgaagtggttgctgttctcatccatacacaa 2220
Db 2161 gataatcacaaagtggttttaatggtgaagtggttgctgttctcatccatacacaa 2220
QY 2221 gatccttatcttgacaatatcttctgtcctgtgtcaatgtcctgagtaacaacctgttccca 2280
Db 2221 gatccttatcttgacaatatcttctgtcctgtgtcaatgtcctgagtaacaacctgttccca 2280
QY 2281 gtccgaattggggacatgacgaatcatgynaataaacttagaaaatgaaatggaagcagt 2340
Db 2281 gtccgaattggggacatgacgaatcatgynaataaacttagaaaatgaaatggaagcagt 2340
QY 2341 attcagcactgaaatcatcccaagctgynaaccagtygtccgatttcttcaatcttctgcta 2400
Db 2341 attcagcactgaaatcatcccaagctgynaaccagtygtccgatttcttcaatcttctgcta 2400
QY 2401 gataaaactgatactttagttattagaacctcctgtcatattgtgccaatagttaaccta 2460
Db 2401 gataaaactgatactttagttattagaacctcctgtcatattgtgccaatagttaaccta 2460
QY 2461 ggtcaagcatcttttgaagccaatgagcatcaatatataatcgacttcaaaaaacttgaa 2520
Db 2461 ggtcaagcatcttttgaagccaatgagcatcaatatataatcgacttcaaaaaacttgaa 2520
QY 2521 ggaataatcatgaccagcatgycagaaaacagccttctgtcatcatatataltatgttttc 2580
Db 2521 ggaataatcatgaccagcatgycagaaaacagccttctgtcatcatatataltatgttttc 2580
QY 2581 cgcctaccaaatacttaaccctaattcatcatcaaccaaggtcctgggggttgggagatca 2640
Db 2581 cgcctaccaaatacttaaccctaattcatcatcaaccaaggtcctgggggttgggagatca 2640
QY 2641 gtgcattatgccaacatggttagatctgcggtgagaccctgcaagccttaattaaatcgt 2700
Db 2641 gtgcattatgccaacatggttagatctgcggtgagaccctgcaagccttaattaaatcgt 2700

Db 2641 gtgcattatgccaacatggttagatctgcggtgagaccctgcaagccttaattaaatcgt 2700
QY 2701 tctcgaagccttagtaataagcaatccaagatatatctgggaactccaagtcacacagatgat 2760
Db 2701 tctcgaagccttagtaataagcaatccaagatatatctgggaactccaagtcacacagatgat 2760
QY 2761 gaagtcgatcaatcatcgggagtaagggttagatcgctccaalttctgggttaaacact 2820
Db 2761 gaagtcgatcaatcatcgggagtaagggttagatcgctccaalttctgggttaaacact 2820
QY 2821 ggtggtccaaaagctgccccatggggatccaaaccccaagtccaagtgcagaatccaacacag 2880
Db 2821 ggtggtccaaaagctgccccatggggatccaaaccccaagtccaagtgcagaatccaacacag 2880
QY 2881 gctatggatcgaaagttgtaatcgtatgtcttcgcaacacagagacgtcaagtttcttacaa 2940
Db 2881 gctatggatcgaaagttgtaatcgtatgtcttcgcaacacagagacgtcaagtttcttacaa 2940
QY 2941 acattaaagggaagcgttaccaactaaaaagccttttcaaggagagctggttcagatg 3000
Db 2941 acattaaagggaagcgttaccaactaaaaagccttttcaaggagagctggttcagatg 3000
QY 3001 gtgtttgcagtygcagcgcttcgggaatcagcttgcacaacagcctgttcttttgaa 3060
Db 3001 gtgtttgcagtygcagcgcttcgggaatcagcttgcacaacagcctgttcttttgaa 3060
QY 3061 ttaatggtaaaagcagatggtgcacacattatactttaatgataaacttgaaggtccaag 3120
Db 3061 ttaatggtaaaagcagatggtgcacacattatactttaatgataaacttgaaggtccaag 3120
QY 3121 aaaaagtcgtttccagaacggtttcatgtagatgacatgcaagctctgtccaagcagattgct 3180
Db 3121 aaaaagtcgtttccagaacggtttcatgtagatgacatgcaagctctgtccaagcagattgct 3180
QY 3181 agtgataatglttccaagattccaagaaggacacagaagaatggttgaagactcaatacaagc 3240
Db 3181 agtgataatglttccaagattccaagaaggacacagaagaatggttgaagactcaatacaagc 3240
QY 3241 ctgcatcttctccaatgatcgtgtgtctgttatgagacagagattgttttagcctt 3300
Db 3241 ctgcatcttctccaatgatcgtgtgtctgttatgagacagagattgttttagcctt 3300
QY 3301 ataaagctcgtctataaacaggtgtcttccaagcttactcatccaagatccaagtgct 3360
Db 3301 ataaagctcgtctataaacaggtgtcttccaagcttactcatccaagatccaagtgct 3360
QY 3361 ctggtgtccttgagcgtgatttttcaagaaatcatctgcagtcatagacactatgtttaca 3420
Db 3361 ctggtgtccttgagcgtgatttttcaagaaatcatctgcagtcatagacactatgtttaca 3420
QY 3421 ttaacttaaccctgcagccttaacttaactccaactgcacatccaactctgtttcttct 3480
Db 3421 ttaacttaaccctgcagccttaacttaactccaactgcacatccaactctgtttcttct 3480
QY 3481 gcaacatctcagaggtcttgatatttctcagaaatgtaacaagccaagaatgcaaatatg 3540
Db 3481 gcaacatctcagaggtcttgatatttctcagaaatgtaacaagccaagaatgcaaatatg 3540
QY 3541 ttggaattatccgtgccttccggccaacagcatattatggcaggaactgtgttaacagag 3600
Db 3541 ttggaattatccgtgccttccggccaacagcatattatggcaggaactgtgttaacagag 3600
QY 3601 ctggctgtcattttagaccctgatgctgaaggaactgttttgataagaagaatgcac 3660
Db 3601 ctggctgtcattttagaccctgatgctgaaggaactgttttgataagaagaatgcac 3660
QY 3661 aatatgtaacacaatttactcttcagtcacgaactcagaccgcggtactctgaacctcag 3720
Db 3661 aatatgtaacacaatttactcttcagtcacgaactcagaccgcggtactctgaacctcag 3720
QY 3721 ataaagctcgaagtyggccatgttgtatactacactctgatgtgtatatacatgnaaaactgta 3780
Db 3721 ataaagctcgaagtyggccatgttgtatactacactctgatgtgtatatacatgnaaaactgta 3780

QY 3781 cctcagctgtatgattttacagaacccacaatcaacgaggaagaccaatltgtatagcc 3840
|||||
Db 3781 cctcagctgtatgattttacagaacccacaatcaacgaggaagaccaatltgtatagcc 3840
QY 3841 actgatgatatagaaagtgagagcggaagtatgataagccagaccglttgcatgycaatc 3900
|||||
Db 3841 actgatgatatagaaagtgagagcggaagtatgataagccagaccglttgcatgycaatc 3900
QY 3901 gcaaggacatcggctccctcaactacaaggccttgcaggttccctcctcaactgcatacagagt 3960
|||||
Db 3901 gcaaggacatcggctccctcaactacaaggccttgcaggttccctcctcaactgcatacagagt 3960
QY 3961 ggcaggcaacacacactaccctttcagcagaatacaagtcgaagccttttgatctgttactt 4020
|||||
Db 3961 ggcaggcaacacacactaccctttcagcagaatacaagtcgaagccttttgatctgttactt 4020
QY 4021 tgggttctcaaaaatgcagatgaaacagttctacagaagtgtttacagatctctcagtc 4080
|||||
Db 4021 tgggttctcaaaaatgcagatgaaacagttctacagaagtgtttacagatctctcagtc 4080
QY 4081 ttgcagctaaaccggcgtattagatctgtcttatactctgtgtctcttgccttgagataaa 4140
|||||
Db 4081 ttgcagctaaaccggcgtattagatctgtcttatactctgtgtctcttgccttgagataaa 4140
QY 4141 gggaaaaaagtgtttgaaacgaatgaatagcttgaacctttaaagaaatcaaaagacatgaa 4200
|||||
Db 4141 gggaaaaaagtgtttgaaacgaatgaatagcttgaacctttaaagaaatcaaaagacatgaa 4200
QY 4201 gcaaaagcttgaagaagctattcttgggagcataggtgccaggcaagaatggtlacgycga 4260
|||||
Db 4201 gcaaaagcttgaagaagctattcttgggagcataggtgccaggcaagaatggtlacgycga 4260
QY 4261 agccgagagacagctcgagagaagccactctggaagtgccttggaaagtcagaanaatttg 4320
|||||
Db 4261 agccgagagacagctcgagagaagccactctggaagtgccttggaaagtcagaanaatttg 4320
QY 4321 aggtggaggaagaatatgactcacttgcgtcaaaaacacagagaagcttgacaaatcaga 4380
|||||
Db 4321 aggtggaggaagaatatgactcacttgcgtcaaaaacacagagaagcttgacaaatcaga 4380
QY 4381 gcagagatcgaacacgaaagcactgtattgtgaaaccttgctacagaagcaaaacctaatc 4440
|||||
Db 4381 gcagagatcgaacacgaaagcactgtattgtgaaaccttgctacagaagcaaaacctaatc 4440
QY 4441 attctagatacattagagatgtgtctcagaacggttctctgaacggaatccaagaagagc 4500
|||||
Db 4441 attctagatacattagagatgtgtctcagaacggttctctgaacggaatccaagaagagc 4500
QY 4501 attcttggtaggtgtctaaaagtgtactactacacagcatgycctgttaaccaaatgtgcagt 4560
|||||
Db 4501 attcttggtaggtgtctaaaagtgtactactacacagcatgycctgttaaccaaatgtgcagt 4560
QY 4561 tatctacaacactgttttgcacacagagagccttgtttcaaagtctcctgaactctta 4620
|||||
Db 4561 tatctacaacactgttttgcacacagagagccttgtttcaaagtctcctgaactctta 4620
QY 4621 ttgtgaagaagagacagagcagtggtgtctgatttatgtcctcagagcttctccgacactgtagc 4680
|||||
Db 4621 ttgtgaagaagagacagagcagtggtgtctgatttatgtcctcagagcttctccgacactgtagc 4680
QY 4681 agtagcatcgggtacaaatacgggtcacaccccccaagtgcctcccttaacctaatlgaggcaa 4740
|||||
Db 4681 agtagcatcgggtacaaatacgggtcacaccccccaagtgcctcccttaacctaatlgaggcaa 4740
QY 4741 aaacttgagattgggaataaacttgcagggtttaaatagcaggtaccaatgtcactatacc 4800
|||||
Db 4741 aaacttgagattgggaataaacttgcagggtttaaatagcaggtaccaatgtcactatacc 4800
QY 4801 tccttggtaggcacatctcagaattttaatgaagaattccttaagacglttctctaagact 4860
|||||
Db 4801 tccttggtaggcacatctcagaattttaatgaagaattccttaagacglttctctaagact 4860

QY 4861 atatgacatatgctgaaagaagatctggaattgaggaaaaacaacattctcgtacaggtc 4920
|||||
Db 4861 atatgacatatgctgaaagaagatctggaattgaggaaaaacaacattctcgtacaggtc 4920
QY 4921 caggatctggtttccaatctccatatgatctcttctgatactgtgaaaaatgaagaacac 4980
|||||
Db 4921 caggatctggtttccaatctccatatgatctcttctgatactgtgaaaaatgaagaacac 4980
QY 4981 caggagatccctgtaaatgttgattgatatctaattgtacagaatttgcgaaggttaccagacc 5040
|||||
Db 4981 caggagatccctgtaaatgttgattgatatctaattgtacagaatttgcgaaggttaccagacc 5040
QY 5041 tctcagaagcgattgcacctgtgttgagaacaatggcaggcaagcacttcagaacggaacat 5100
|||||
Db 5041 tctcagaagcgattgcacctgtgtgtgagaacaatggcaggcaagcacttcagaacggaacat 5100
QY 5101 catgctgaagctgcacagtggtctagttccactcagcagcaactgtgtgtaaatattgagc 5160
|||||
Db 5101 catgctgaagctgcacagtggtctagttccactcagcagcaactgtgtgtaaatattgagc 5160
QY 5161 atgctggagacgggaatatctctcgtgtggatgtgtaacattccagaatatctcatct 5220
|||||
Db 5161 atgctggagacgggaatatctctcgtgtggatgtgtaacattccagaatatctcatct 5220
QY 5221 aatgttttagaagaatctgcggtctcagatgatgtgtatctccaagatgaagaaggtatc 5280
|||||
Db 5221 aatgttttagaagaatctgcggtctcagatgatgtgtatctccaagatgaagaaggtatc 5280
QY 5281 tgcctcggaaaaatcacttactagtcagagccttgtggattactgtgaacaaagcagctgct 5340
|||||
Db 5281 tgcctcggaaaaatcacttactagtcagagccttgtggattactgtgaacaaagcagctgct 5340
QY 5341 tccttctctatgctgctgcaltgtatgaagcagtttaatgaagtttaacaaagtactatctct 5400
|||||
Db 5341 tccttctctatgctgctgcaltgtatgaagcagtttaatgaagtttaacaaagtactatctct 5400
QY 5401 attcatgaagcttaatcgggagtgcgaagaactataccacaatttcagtgtaaacttcaagaa 5460
|||||
Db 5401 attcatgaagcttaatcgggagtgcgaagaactataccacaatttcagtgtaaacttcaagaa 5460
QY 5461 gcatcagcaaaaatgttcatcagagtlactgcgtggagcggatgttggcacctatttc 5520
|||||
Db 5461 gcatcagcaaaaatgttcatcagagtlactgcgtggagcggatgttggcacctatttc 5520
QY 5521 cgtgttggtttttatggaaccaagtctcgggatttggatgaaacaagaatttgtttacaag 5580
|||||
Db 5521 cgtgttggtttttatggaaccaagtctcgggatttggatgaaacaagaatttgtttacaag 5580
QY 5581 gagcctgcataataccaaaacttgcagagatactcaacagattggaaggattttaacggagaa 5640
|||||
Db 5581 gagcctgcataataccaaaacttgcagagatactcaacagattggaaggattttaacggagaa 5640
QY 5641 agatttggagagagtgtgttgaagtaatcaaaagactctaactcctgtagacaagtgtaaa 5700
|||||
Db 5641 agatttggagagagtgtgttgaagtaatcaaaagactctaactcctgtagacaagtgtaaa 5700
QY 5701 ttagatccctaacaaggcatataatcagattacctaagtgtgagccaactttgacacatat 5760
|||||
Db 5701 ttagatccctaacaaggcatataatcagattacctaagtgtgagccaactttgacacatat 5760
QY 5761 gagaagaagacagaaatcacctatttcgacaanaaatlacaaatctcgtcgatltcatgtac 5820
|||||
Db 5761 gagaagaagacagaaatcacctatttcgacaanaaatlacaaatctcgtcgatltcatgtac 5820
QY 5821 tgtacaccccttactttagatggtccgttgcctatggtgggaacttcaagaacttcaaaagg 5880
|||||
Db 5821 tgtacaccccttactttagatggtccgttgcctatggtgggaacttcaagaacttcaaaagg 5880
QY 5881 aagaacatctgaactaagtcctcatgcttctctataatlaaaacaaggttcaatgttcaact 5940
|||||
Db 5881 aagaacatctgaactaagtcctcatgcttctctataatlaaaacaaggttcaatgttcaact 5940
QY 5941 cataaagaagatcatcttaacaccaaatltgaagltgtctatltgaggacatgcagaaaaag 6000

Db 5941 cataaagaagatcatcttaacaccaattgaagttgctatctagacatgcagaaaaag 6000
QY 6001 acacaggagttggcatttgcaacacatcagatcccgagacccccaaaatgtctcagatg 6060
Db 6001 acacaggagttggcatttgcaacacatcagatcccgagacccccaaaatgtctcagatg 6060
QY 6061 gtactccaaggatctgtagggcacccacagtgatcaggggcctttggaaagttggccaggtt 6120
Db 6061 gtactccaaggatctgtagggcacccacagtgatcaggggcctttggaaagttggccaggtt 6120
QY 6121 ttctgtctgaaatacctagtgaccccaagctcttcagacatcataataaactgcgactc 6180
Db 6121 ttctgtctgaaatacctagtgaccccaagctcttcagacatcataataaactgcgactc 6180
QY 6181 tgctttaaagatttactaaaagtgltgaagatgccttaagaaaaaataagagcttaatt 6240
Db 6181 tgctttaaagatttactaaaagtgltgaagatgccttaagaaaaaataagagcttaatt 6240
QY 6241 gggccggttcaaaaaggatcaaaagggaattggggaaactatctcgcccttaagaggc 6300
Db 6241 gggccggttcaaaaaggatcaaaagggaattggggaaactatctcgcccttaagaggc 6300
QY 6301 cctacagccctagatcacagaagaagtcocctcagttatccaagccagtatctgtctccct 6360
Db 6301 cctacagccctagatcacagaagaagtcocctcagttatccaagccagtatctgtctccct 6360
QY 6361 gccacaagagattccttcagtcgaatgagcttcgcgcaaaatggaatcctctaactgaatgca 6420
Db 6361 gccacaagagattccttcagtcgaatgagcttcgcgcaaaatggaatcctctaactgaatgca 6420
QY 6421 ctgttttattcatctgcgaaaggccatgtaattcaacatcgagtgtaaaagatctattg 6480
Db 6421 ctgttttattcatctgcgaaaggccatgtaattcaacatcgagtgtaaaagatctattg 6480
QY 6481 gaaccaacatggaatggaatctctggaattatattcatcattgaaagatgcagtcgccaag 6540
Db 6481 gaaccaacatggaatggaatctctggaattatattcatcattgaaagatgcagtcgccaag 6540
QY 6541 aaatatcaaatgtagattgtttaacgcttgagaatcagtggttctctaattgttcgg 6600
Db 6541 aaatatcaaatgtagattgtttaacgcttgagaatcagtggttctctaattgttcgg 6600
QY 6601 gtaacaagctgtatctttaaagacattttaatgactccaaggtacacatatatactacc 6660
Db 6601 gtaacaagctgtatctttaaagacattttaatgactccaaggtacacatatatactacc 6660
QY 6661 attattataccatagcctaaggttaaaaaattatctcaactttaagttcgtatttttaatt 6720
Db 6661 attattataccatagcctaaggttaaaaaattatctcaactttaagttcgtatttttaatt 6720
QY 6721 tataccaccatttatagattcatttggaccattttaaatgtagtaatgcttatttaa 6780
Db 6721 tataccaccatttatagattcatttggaccattttaaatgtagtaatgcttatttaa 6780
QY 6781 aggtactaaaaaataatgtgaatgtttaccctgtgcgcgccaagggcctc 6828
Db 6781 aggtactaaaaaataatgtgaatgtttaccctgtgcgcgccaagggcctc 6828

RESULT 2
AAH41911
ID AAH41911 standard; cDNA; 4143 BP.
XX
AC AAH41911;
XX
DT 30-AUG-2001 (first entry)
XX
DE Preliminary human CLASP-3 cDNA sequence Fig 1.
XX
KW Human; cadherin-like asymmetry protein; CLASP; CLASP-3; immune response;
cell surface molecule; transmembrane protein; immunosuppressive; vaccine;
antiinflammatory; antiarthritic; antianaemic; dermatological; uropathic;
KW

KW ophthalmological; antirheumatic; nephrotropic; antithyroid; antidiabetic;
KW neuroprotective; antiasthmatic; antibacterial; antisense therapy;
KW gene therapy; chromosome 1p31.1; ss.
OS Homo sapiens.
XX
PN WO200142297-A2.
XX
PD 14-JUN-2001.
XX
PF 13-DEC-2000; 2000WO-US34171.
XX
PR 13-DEC-1999; 99US-0170453.
PR 14-JAN-2000; 2000US-0176195.
PR 14-FEB-2000; 2000US-0182296.
PR 11-APR-2000; 2000US-0196267.
PR 11-APR-2000; 2000US-0196460.
PR 11-APR-2000; 2000US-0196527.
PR 11-APR-2000; 2000US-0196528.
PR 11-APR-2000; 2000US-0547276.
PR 13-OCT-2000; 2000US-0240503.
PR 13-OCT-2000; 2000US-0240508.
XX
PA (ARBO-) ARBOR VITA CORP.
XX
PI Lu P, Garman JD, Candia AF;
XX
DR WPI; 2001-375003/39.
DR P-PSDB; AAB99493.
XX
PT Novel isolated cadherin-like asymmetry protein (CLASP) -3, useful for
PT useful in preventing or treating a CLASP-3-mediated disease e.g.
PT autoimmune disease -
XX
PS Example 5; Fig 1; 189pp; English.
XX
CC The present invention describes an isolated polypeptide (I) comprising a
CC nucleotide (nt) sequence (SIa) that has at least 90% identity to the
CC sequence given in AAH41934 and is immunologically cross-reactive with
CC the derived amino acid (aa) sequence (SIb) given in AAB99495 or shares a
CC biological function with native CLASP-3. (I) has immunosuppressive,
CC antiinflammatory, antiarthritic, antianaemic, dermatological, uropathic,
CC ophthalmological, antirheumatic, nephrotropic, antithyroid, antidiabetic,
CC neuroprotective, antiasthmatic and antibacterial activities, and can be
CC used in antisense therapy, vaccine production and gene therapy. CLASP-3
CC related sequences can be used in preventing or treating a CLASP-3
CC mediated disease, preferably an autoimmune disease by inhibiting an
CC immune response. The autoimmune disease is caused or exacerbated by
CC increased activity of T helper 1 lymphocytes (TH1). AAH41893 to AAH41953
CC and AAB99491 to AAB99507 represent sequences which are used in the
CC exemplification of the present invention. CLASP-3 is localised in the
CC chromosome location 1p31.1.
XX
SQ Sequence 4143 BP; 1251 A; 829 C; 890 G; 1173 T; 0 other;

Query Match 60.4%; Score 4123; DB 22; Length 4143;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 4136; Conservative 0; Mismatches 0; Indels 3; Gaps 1;
QY 2693 taaatcgttctcgaagccttagtaatagcaatccagatatatctggactcccacgctcac 2752
Db 5 taaatcgttctcgaagccttagtaatagcaatccagatatatctggactcccacgctcac 64
QY 2753 cagatgatgaagttcgatcaatcatcgcgggagtaagggttagatcgcctcaattcctggg 2812
Db 65 cagatgatgaagttcgatcaatcatcgcgggagtaagggttagatcgcctcaattcctggg 124
QY 2813 ttaacactggtgtgccaaaagctgcccacatggygatccaacccacagtccaagtgcagaat 2872
Db 125 ttaacactggtgtgccaaaagctgcccacatggygatccaacccacagtccaagtgcagaat 184
QY 2873 caacacaggtatgatcgaaagttgtaatcgtatgtcttcgcacacagagaagcgtcaagtt 2932

|||||
Db 185 caacacagcgtacgtgcatcgaaagtgtgaatcgtctgcacacagagacgtcaagtt 244
QY 2933 tcttaacaacatlaaacggygacgcttaccaactaaaaagcttttccagagagctgctt 2992
Db 245 tcttaacaacatlaaacggygacgcttaccaactaaaaagcttttccagagagctgctt 304
QY 2993 tgcagtggtgtgttgcagtgagtcgagcgttcgggaatcagctttgcacaagcctgttct 3052
Db 305 tgcagtggtgtgttgcagtgagtcgagcgttcgggaatcagctttgcacaagcctgttct 364
QY 3053 ttttgaattaatggtlaaagagcatgtgtgacccattatactttaatgataaacttgag 3112
Db 365 ttttgaattaatggtlaaagagcatgtgtgacccattatactttaatgataaacttgag 424
QY 3113 ctccaaggaanaagtcgtttccagaaacgtttcatgtgatgacatgtgcagctctgtcagca 3172
Db 425 ctccaaggaanaagtcgtttccagaaacgtttcatgtgatgacatgtgcagctctgtcagca 484
QY 3173 cgattgtcagtgtatagtttccacgatttccagaagacacagaaatggttgagagactca 3232
Db 485 cgattgtcagtgtatagtttccacgatttccagaagacacagaaatggttgagagactca 544
QY 3233 atacaagccttgcatcttctcgaatgacgtgtgtctgttlatgacagagagattgttt 3292
Db 545 atacaagccttgcatcttctcgaatgacgtgtgtctgttlatgacagagagattgttt 604
QY 3293 ttagccttataaagtcctgtcataaagagtgcttcaaaagcttactcatlaccgaatc 3352
Db 605 ttagccttataaagtcctgtcataaagagtgcttcaaaagcttactcatlaccgaatc 664
QY 3353 ccaagtgtctgtgtcctttagagctggaatttctacgaatcatctgcagtcagtcagcact 3412
Db 665 ccaagtgtctgtgtcctttagagctggaatttctacgaatcatctgcagtcagtcagcact 724
QY 3413 atgttacatlaaacctaccctgcagcttacttactccaacctgcacctccatcaccttctg 3472
Db 725 atgttacatlaaacctaccctgcagcttacttactccaacctgcacctccatcaccttctg 784
QY 3473 ttctcttcgcaacatctcagagttctgtgattttctacgaatgttacaagaccaaaagattg 3532
Db 785 ttctcttcgcaacatctcagagttctgtgattttctacgaatgttacaagaccaaaagattg 844
QY 3533 caaatatgtttgaattatccgtgccttccggccaacagcatlatttggcaggactgtgt 3592
Db 845 caaatatgtttgaattatccgtgccttccggccaacagcatlatttggcaggactgtgt 904
QY 3593 taacagagctgtgctgcatltagaacctgtagtctgaaagactgtttggaattgcataaga 3652
Db 905 taacagagctgtgctgcatltagaacctgtagtctgaaagactgtttggaattgcataaga 964
QY 3653 aagtcacataatgtgtacacaatttactctccagtcacgactcagacccgcggtactctg 3712
Db 965 aagtcacataatgtgtgtacacaatttactctccagtcacgactcagacccgcggtactctg 1024
QY 3713 accctcagataaaggctcgagtggtgcatgtgtatctacacctgtgattgtattatcatg 3772
Db 1025 accctcagataaaggctcgagtggtgcatgtgtatctacacctgtgattgtattatcatg 1084
QY 3773 aaactgtacctcagctgtatgtattttacagaaaactcacaaatcaacgaggaagaccaatt 3832
Db 1085 aaactgtacctcagctgtatgtattttacagaaaactcacaaatcaacgaggaagaccaatt 1144
QY 3833 gtagagcactgtatgtattatgaaaagtgaagcggaagtatgataagccagacggttgcca 3892
Db 1145 gtagagcactgtatgtattatgaaaagtgaagcggaagtatgataagccagacggttgcca 1204
QY 3893 tggcaatcgcaaggacatcggtccctcaactaaagcctgscagttcctcctcacgt 3952
Db 1205 tggcaatcgcaaggacatcggtccctcaactaaagcctgscagttcctcctcacgt 1264
QY 3953 caacgagtggcaggcaacacactacctttcagcgaatcaagtgcgaagcctttgtatct 4012
|||||

Db 1265 caacgagtggcaggcaacacactaccttttcagcgaatcaagtcgaagcctttgtatct 1324
QY 4013 gtcactcttggttctcaaaaaatgcagatgtgaacagttctacagaagtgtttacagatc 4072
Db 1325 gtcactcttggttctcaaaaaatgcagatgtgaacagttctacagaagtgtttacagatc 1384
QY 4073 tctcagctcttgcaagctaaacccggtcattagatctgtcttatactctgtgtcttgccttg 4132
Db 1385 tctcagctcttgcaagctaaacccggtcattagatctgtcttatactctgtgtcttgccttg 1444
QY 4133 agtataaagggaanaagtggtttgaaacgaatgatatagcttgaccttlaagaaatcaaaag 4192
Db 1445 agtataaagggaanaagtggtttgaaacgaatgatatagcttgaccttlaagaaatcaaaag 1504
QY 4193 acatgagagcaaaagcttgaagaagctatcttgggagcataggtgccaggcaagaatgg 4252
Db 1505 acatgagagcaaaagcttgaagaagctatcttgggagcataggtgccaggcaagaatgg 1564
QY 4253 tacggcgaaagccgagacagctcgagagaagcccatctgaaagtgcctttggaagtlcaag 4312
Db 1565 tacggcgaaagccgagacagctcgagagaagcccatctgaaagtgcctttggaagtlcaag 1624
QY 4313 aaaattgagtgtaggaagaatatgtactcacctggcggtcaaaaacacagagaagcttgaca 4372
Db 1625 aaaattgagtgtaggaagaatatgtactcacctggcggtcaaaaacacagagaagcttgaca 1684
QY 4373 aatcaagagcagagatgtgaacacggaagcactgtattgtgaaacctgtgctacagaagcaa 4432
Db 1685 aatcaagagcagagatgtgaacacggaagcactgtattgtgaaacctgtgctacagaagcaa 1744
QY 4433 acctaatcattttagatcacattagagattgttctcagaacggttctctgtaacggaatcca 4492
Db 1745 acctaatcattttagatcacattagagattgttctcagaacggttctctgtaacggaatcca 1804
QY 4493 aagagagcattctgtgtgagtgcttaaaagtgtctaactacacagcatgtgcctgtlaaccaaa 4552
Db 1805 aagagagcattctgtgtgagtgcttaaaagtgtctaactacacagcatgtgcctgtlaaccaaa 1864
QY 4553 gtgcagttatctacaacactgttltgtcacacagagagccttgttccaagtttccctg 4612
Db 1865 gtgcagttatctacaacactgttltgtcacacagagagccttgttccaagtttccctg 1924
QY 4613 aactctatttgaagaagagacagagcagtggtgtgattttagcctcaggtcttccgac 4672
Db 1925 aactctatttgaagaagagacagagcagtggtgtgattttagcctcaggtcttccgac 1984
QY 4673 actgtagcagtagcatcgtgtacaatacaggttcaacccccagtgctcccttaccactaa 4732
Db 1985 actgtagcagtagcatcgtgtacaatacaggttcaacccccagtgctcccttaccactaa 2044
QY 4733 tgaaggcaaaaactttgagattgggaataactttgcagggtttaaagtgcaggttaccaatgt 4792
Db 2045 tgaaggcaaaaactttgagattgggaataactttgcagggtttaaagtgcaggttaccaatgt 2104
QY 4793 caatatcctcctgtgtgycacacatctcagaatlttaatgaaaatctttaaagacgttctc 4852
Db 2105 caatatcctcctgtgtgycacacatctcagaatlttaatgaaaatctttaaagacgttctc 2164
QY 4853 taaagaactatatgacatatgtctgaagaagatctgaaattgagggaaacaacatltccctg 4912
Db 2165 taaagaactatatgacatatgtctgaagaagatctgaaattgagggaaacaacatltccctg 2224
QY 4913 atcaggtccagatctggtttcaactctccatacgattcttctgatactgtgaaatgta 4972
Db 2225 atcaggtccagatctggtttcaactctccatacgattcttctgatactgtgaaatgta 2284
QY 4973 aggaacacccaggaagatcctgaaatgtttgatgtactaatgttacagaattgccaaggtt 5032
Db 2285 aggaacacccaggaagatcctgaaatgtttgatgtactaatgttacagaattgccaaggtt 2344
QY 5033 accagacctctccaga---gcgattgacctgtgtgcagaacatgycgagcaagcactcag 5089
Db 2345 accagacctctccagatctgcattgacctgtgtgcagaacatgycgagcaagcactcag 2404
|||||

QY 5090 aacgaagcaatcatgctgaagctgcacagtgctctagtcacctcagcagcactgtgtctg 5149
|||||
Db 2405 aacgaagcaatcatgctgaagctgcacagtgctctagtcacctcagcagcactgtgtctg 2464
QY 5150 aatatgtgacatgctggaagccggaatatctccctgtggagtgtgtacaattcaga 5209
|||||
Db 2465 aatatgtgacatgctggaagccggaatatctccctgtggagtgtgtacaattcaga 2524
QY 5210 atattcatctaatgttttagaagaatctgcggtctcagatgtgtgtatctccagatg 5269
|||||
Db 2525 atattcatctaatgttttagaagaatctgcggtctcagatgtgtgtatctccagatg 2584
QY 5270 aagaaggtatctgctctggaaaaatacttactgagtcaggaactgtgagatactggaac 5329
|||||
Db 2585 aagaaggtatctgctctggaaaaatacttactgagtcaggaactgtgagatactggaac 2644
QY 5330 aagaagctgctctcctctctatgctggaatgtatgaagcagttaatgaagttacaag 5389
|||||
Db 2645 aagaagctgctctcctctctatgctggaatgtatgaagcagttaatgaagttacaag 2704
QY 5390 tactatctcattatcatgaagctaalcgagatgcagaagaaactataccacaattcatgta 5449
|||||
Db 2705 tactatctcattatcatgaagctaalcgagatgcagaagaaactataccacaattcatgta 2764
QY 5450 aactcaagaagcattcagcaaaaattgttcacagagctggtggtgagcggaattgt 5509
|||||
Db 2765 aactcaagaagcattcagcaaaaattgttcacagagctggtggtgagcggaattgt 2824
QY 5510 gcacctatttcgtgtgtgttttatggaaccaagttcggggaatttgatgaacaagaat 5569
|||||
Db 2825 gcacctatttcgtgtgtgttttatggaaccaagttcggggaatttgatgaacaagaat 2884
QY 5570 ttgtttacaaggaagcctgcataaaccacttgcagagatatctcacagattggaaggat 5629
|||||
Db 2885 ttgtttacaaggaagcctgcataaaccacttgcagagatatctcacagattggaaggat 2944
QY 5630 tttaacgagaagaatttggagaagatgtgttgaagtaatcaaaagactctaactctgtag 5689
|||||
Db 2945 tttaacgagaagaatttggagaagatgtgttgaagtaatcaaaagactctaactctgtag 3004
QY 5690 acaagtgtaaattagatccctaacaaggcataatatcagattacctaattgtgagccatact 5749
|||||
Db 3005 acaagtgtaaattagatccctaacaaggcataatatcagattacctaattgtgagccatact 3064
QY 5750 ttgacacatatgagatgaaagacagaatcacactatttcgacaaaaaattacaatcttcgtc 5809
|||||
Db 3065 ttgacacatatgagatgaaagacagaatcacactatttcgacaaaaaattacaatcttcgtc 3124
QY 5810 gattcaatgtactgtacaccccttactttagatgcccgtgcccatgvggaacttcatagaac 5869
|||||
Db 3125 gattcaatgtactgtacaccccttactttagatgcccgtgcccatgvggaacttcatagaac 3184
QY 5870 aattcaaaaaggaagaccattctgaactacgtctcatgaccttccctatatataaacaaggg 5929
|||||
Db 3185 aattcaaaaaggaagaccattctgaactacgtctcatgaccttccctatatataaacaaggg 3244
QY 5930 tcaatgtcactcataaagaagagatcatcttaacccaattgaagttgctatgagaca 5989
|||||
Db 3245 tcaatgtcactcataaagaagagatcatcttaacccaattgaagttgctatgagaca 3304
QY 5990 tgcgaaaaaagacacagagttggtgcaatttgcaacacatcaggtatcccgcacagcccaaaa 6049
|||||
Db 3305 tgcgaaaaaagacacagagttggtgcaatttgcaacacatcaggtatcccgcacagcccaaaa 3364
QY 6050 tgcctcagatgtgtactccagggatctgtlaggcacacacagtgaaatcaggggaccttggaag 6109
|||||
Db 3365 tgcctcagatgtgtactccagggatctgtlaggcacacacagtgaaatcaggggaccttggaag 3424
QY 6110 ttgcccaaggttttctgtctgaaatacctagtgacccaaggctcttcagacatacata 6169
|||||
Db 3425 ttgcccaaggttttctgtctgaaatacctagtgacccaaggctcttcagacatacata 3484

QY 6170 aactgcgactctgctttaagaatttacttaaaaggtgtgaagatgaccttaagaaaaata 6229
|||||
Db 3485 aactgcgactctgctttaagaatttacttaaaaggtgtgaagatgaccttaagaaaaata 3544
QY 6230 agagcttaattggccggtccaagaagaglatcaaaaggaattggggaactatcttcgc 6289
|||||
Db 3545 agagcttaattggccggtccaagaagaglatcaaaaggaattggggaactatcttcgc 3604
QY 6290 ctlaaagagccctacagccctagatcacagaagagtcctcagttatccaagccagta 6349
|||||
Db 3605 ctlaaagagccctacagccctagatcacagaagagtcctcagttatccaagccagta 3664
QY 6350 gctgtccctgcacagagatccctcagtcgaatgagcttcgcaaatggtatctcta 6409
|||||
Db 3665 gctgtccctgcacagagatccctcagtcgaatgagcttcgcaaatggtatctcta 3724
QY 6410 aactgaatgcaactgttttatcatctgcaaaagagccatgtrattcaacatcgagtga 6469
|||||
Db 3725 aactgaatgcaactgttttatcatctgcaaaagagccatgtrattcaacatcgagtga 3784
QY 6470 aagatctatggaaaccacaatggaatggaattctggaatatattcatgtgaagaatg 6529
|||||
Db 3785 aagatctatggaaaccacaatggaatggaattctggaatatattcatgtgaagaatg 3844
QY 6530 cagtgccaagaaaatatcaaatgtagaattgtlaacgcttgagaatcatggtatggtt 6589
|||||
Db 3845 cagtgccaagaaaatatcaaatgtagaattgtlaacgcttgagaatcatggtatggtt 3904
QY 6590 ctaatgtcgggtaacaagctgtatcttttaagacattttaatgactcaaaaggtacact 6649
|||||
Db 3905 ctaatgtcgggtaacaagctgtatcttttaagacattttaatgactcaaaaggtacact 3964
QY 6650 atacattacattattatataccatagctlaaggtlaaaaattattcaactttaagtcgt 6709
|||||
Db 3965 atacattacattattatataccatagctlaaggtlaaaaattattcaactttaagtcgt 4024
QY 6710 atttttaattatataccacatttagattcattttggaccacattttaatgtagta 6769
|||||
Db 4025 atttttaattatataccacatttagattcattttggaccacattttaatgtagta 4084
QY 6770 gctattttaaaggtactaaaaaatgtgaatgtttacccgtgcgcgcagggccctc 6828
|||||
Db 4085 gctattttaaaggtactaaaaaatgtgaatgtttacccgtgcgcgcagggccctc 4143

RESULT 3
AAH41912
ID AAH41912 standard; cDNA; 4144 BP.
XX
AC AAH41912;
XX
AC 30-AUG-2001 (first entry)
DT
DT 30-AUG-2001 (first entry)
XX
DE Human CLASP-3 cDNA sequence Fig 4.
XX
XX Human; cadherin-like asymmetry protein; CLASP; CLASP-3; immune response;
KW cell surface molecule; transmembrane protein; immunosuppressive; vaccine;
KW antiinflammatory; antiarthritic; antianaemic; dermatological; uropathic;
KW ophthalmological; antirheumatic; nephroretropic; antithyroid; antidiabetic;
KW neuroprotective; antiasthmatic; antibacterial; antisense therapy;
KW gene therapy; chromosome 1p31.1; ss.
XX
OS Homo sapiens.
OS
XX
PN WO200142297-A2.
PN
XX
PD 14-JUN-2001.
PD
XX
PF 13-DEC-2000; 2000WO-US34171.
PF
XX
PR 13-DEC-1999; 99US-0170453.
PR
PR 14-JAN-2000; 2000US-0176195.
PR
PR 14-FEB-2000; 2000US-0182296.

PR 11-APR-2000; 2000US-0196267.
PR 11-APR-2000; 2000US-0196460.
PR 11-APR-2000; 2000US-0196527.
PR 11-APR-2000; 2000US-0196528.
PR 11-APR-2000; 2000US-0547276.
PR 13-OCT-2000; 2000US-0240503.
PR 13-OCT-2000; 2000US-0240508.
XX
PA (ARBO-) ARBOR VITA CORP.
XX
PI Lu P, Garman JD, Candia AF;
XX
DR WPI: 2001-375003/39.
DR P-PSDB; AAB99494.
XX
PT Novel isolated cadherin-like asymmetry protein (CLASP) -3, useful for
PT useful in preventing or treating a CLASP-3-mediated disease e.g.
PT autoimmune disease
XX
PS Disclosure; Fig 4A; 189pp; English.

XX
CC The present invention describes an isolated polypeptide (I) comprising a
CC nucleotide (nt) sequence (S1a) that has at least 90% identity to the
CC sequence given in AAH41934 and is immunologically cross-reactive with
CC the derived amino acid (aa) sequence (S1b) given in AAB99495 or shares a
CC biological function with native CLASP-3. (I) has immunosuppressive,
CC antiinflammatory, antiarthritic, antianaemic, dermatological, uropathic,
CC ophthalmological, antirheumatic, nephrotropic, antithyroid, antidiabetic,
CC neuroprotective, antiasthmatic and antibacterial activities, and can be
CC used in antisense therapy, vaccine production and gene therapy. CLASP-3
CC related sequences can be used in preventing or treating a CLASP-3
CC mediated disease, preferably an autoimmune disease by inhibiting an
CC immune response. The autoimmune disease is caused or exacerbated by
CC increased activity of T helper 1 lymphocytes (TH1). AAH41893 to AAH41953
CC and AAB99491 to AAB99507 represent sequences which are used in the
CC exemplification of the present invention. CLASP-3 is localised in the
CC chromosome location 1p31.1.
XX
SQ Sequence 4144 BP; 1252 A; 829 C; 890 G; 1173 T; 0 other;

Query Match 60.4%; Score 4123; DB 22; Length 4144;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 4136; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

QY 2693 taatcgtcttcgaagccttagtaatagcaatccagatatatctgaggactcccagctcac 2752
Db 6 taatcgtcttcgaagccttagtaatagcaatccagatatatctgaggactcccagctcac 65
QY 2753 cagatgatgaagtcgatcaatcatcgggaagtaagggttagatgcctccaattcctgg 2812
Db 66 cagatgatgaagtcgatcaatcatcgggaagtaagggttagatgcctccaattcctgg 125
QY 2813 ttaacactgtgtgccaaaagctgccccatggggatccaaccccgatccaagtgcagaat 2872
Db 126 ttaacactgtgtgccaaaagctgccccatggggatccaaccccgatccaagtgcagaat 185
QY 2873 caacacaggtatgtagcgaagtgtatacgtatgtcttcgcacacagagagcgtcaagt 2932
Db 186 caacacaggtatgtagcgaagtgtatacgtatgtcttcgcacacagagagcgtcaagt 245
QY 2933 tcttacaacattaaacgggagcgttaccactaaaagctttttcaagaggagctggtc 2992
Db 246 tcttacaacattaaacgggagcgttaccactaaaagctttttcaagaggagctggtc 305
QY 2993 tgcagtggtgttttgcaagtggcagcgttcgggaaatcagcttgcacaacaggcctgttct 3052
Db 306 tgcagtggtgttttgcaagtggcagcgttcgggaaatcagcttgcacaacaggcctgttct 365
QY 3053 ttttgaatlaatggtaaagagcatgtgtcacacattatactttaatgataaactgag 3112
Db 366 ttttgaatlaatggtaaagagcatgtgtcacacattatactttaatgataaactgag 425

QY 3113 ctccaagaaaagtcgtttccagaacggttccatgatgacattgacgctctgtlcagca 3172
Db 426 ctccaagaaaagtcgtttccagaacggttccatgatgacattgacgctctgtlcagca 485
QY 3173 cgattgtagtatagatttcacgatttcagaaggacacagaatatgtgtgagagactca 3232
Db 486 cgattgtagtatagatttcacgatttcagaaggacacagaatatgtgtgagagactca 545
QY 3233 atacaagccttgcatcttctcgaatgatctgtgtctgttatgacagaggattgttt 3292
Db 546 atacaagccttgcatcttctcgaatgatctgtgtctgttatgacagaggattgttt 605
QY 3293 ttagccttataaagtcctgctataaacaagtgctcttcaaaagcttactcattaccgaatc 3352
Db 606 ttagccttataaagtcctgctataaacaagtgctcttcaaaagcttactcattaccgaatc 665
QY 3353 ccaagttctgtgtccttgaggctggattttctaagaaatcatctgcagtcatgagcaat 3412
Db 666 ccaagttctgtgtccttgaggctggattttctaagaaatcatctgcagtcatgagcaat 725
QY 3413 atgttaacattaaacttaccctgcagcttacttaccacctgcacatccacactctc 3472
Db 726 atgttaacattaaacttaccctgcagcttacttaccacctgcacatccacactctc 785
QY 3473 ttctctcgcaacatctcagagttctgattttctacgaatgtlacaagaccaaaagattg 3532
Db 786 ttctctcgcaacatctcagagttctgattttctacgaatgtlacaagaccaaaagattg 845
QY 3533 caaatatgtttgaattatccgtgcctttccgccaacagcattatttggcaggacttgt 3592
Db 846 caaatatgtttgaattatccgtgcctttccgccaacagcattatttggcaggacttgt 905
QY 3593 taacagaagctgctgtcatctttagaccctgatgctggaagagctgtttgattgcataaga 3652
Db 906 taacagaagctgctgtcatctttagaccctgatgctggaagagctgtttgattgcataaga 965
QY 3653 aagtcatacaatagtgtacacaatttactctccagtcacgactcagaccgcgylactctg 3712
Db 966 aagtcatacaatagtgtacacaatttactctccagtcacgactcagaccgcgylactctg 1025
QY 3713 accctcagataaaggctcagagtgccatgtgtlatacctctgattgtattatcatg 3772
Db 1026 accctcagataaaggctcagagtgccatgtgtlatacctctgattgtattatcatg 1085
QY 3773 aaactgtacctcagctgtcatgatttacagaacctcaacatcaacgaggaagccaatt 3832
Db 1086 aaactgtacctcagctgtcatgatttacagaacctcaacatcaacgaggaagccaatt 1145
QY 3833 gtagtagccactgatgatattgaagaagtgaagcggaagtatgataagccagaccgttgcca 3892
Db 1146 gtagtagccactgatgatattgaagaagtgaagcggaagtatgataagccagaccgttgcca 1205
QY 3893 tggcaatcgcaaggacatcgttccctcaactaaacaggcctgycagtttccctctcaagt 3952
Db 1206 tggcaatcgcaaggacatcgttccctcaactaaacaggcctgycagtttccctctcaagt 1265
QY 3953 caacgagtggcaggaacacacactacctttcagcagaatcaagtgcgaagcctttgatct 4012
Db 1266 caacgagtggcaggaacacacactacctttcagcagaatcaagtgcgaagcctttgatct 1325
QY 4013 gtctactttgggttctcaaaaatgcatatgaacagtttctacagaagtgtttacagatc 4072
Db 1326 gtctactttgggttctcaaaaatgcatatgaacagtttctacagaagtgtttacagatc 1385
QY 4073 tctcagttcttcagctaaacccggtattagatctgtcttatctctgtgtcttgccttg 4132
Db 1386 tctcagttcttcagctaaacccggtattagatctgtcttatctctgtgtcttgccttg 1445
QY 4133 agtataaagggaataaaggttttgaacgaatgaatagcttgaccttcaagaaatcaaaag 4192
Db 1446 agtataaagggaataaaggttttgaacgaatgaatagcttgaccttcaagaaatcaaaag 1505
QY 4193 acatgagagcaaaagcttgaagaagctatcttctgggagcataggtgccaggcaagaatgg 4252

Db 1506 acatgagacaaagcttgaagaagctatcttgygagcatagtgccagccaagaatgg 1565
QY 4253 tacgycgaagccgagacagctcgagagaagcccatctggaagtgccttgyaagtcag 4312
Db 1566 tacgycgaagccgagacagctcgagagaagcccatctggaagtgccttgyaagtcag 1625
QY 4313 aaatttgaggtgaggaagaatatgactcactggcgtcacaacacagagaagcttgaca 4372
Db 1626 aaaatttgaggtgaggaagaatatgactcactggcgtcacaacacagagaagcttgaca 1685
QY 4373 aatcaagagcagaagattgaaacaggaactgattgatgaaaacctggctacagaagcaa 4432
Db 1686 aatcaagagcagaagattgaaacaggaactgattgatgaaaacctggctacagaagcaa 1745
QY 4433 accaatcattttaagatacatltagaagattggtltcagaccgtttctgtaacgaaatcca 4492
Db 1746 accaatcattttaagatacatltagaagattggtltcagaccgtttctgtaacgaaatcca 1805
QY 4493 aagagagcatctctgtgtgagtgctaaaagtgtactacacagcatggcctgttaaccaa 4552
Db 1806 aagagagcatctctgtgtgagtgctaaaagtgtactacacagcatggcctgttaaccaa 1865
QY 4553 gtgcagtttatctacaacactgttttgtctacacagagagccttggtttccaaagtttcctg 4612
Db 1866 gtgcagtttatctacaacactgttttgtctacacagagagccttggtttccaaagtttcctg 1925
QY 4613 aactcttatttgaagaagaagacagagcagtggtgtctgattatgtccctcagccttctccgac 4672
Db 1926 aactcttatttgaagaagaagacagagcagtggtgtctgattatgtccctcagccttctccgac 1985
QY 4673 actgtagcagtgacatcgytacaacagtcacaccccagtgccctcccttacctactaa 4732
Db 1986 actgtagcagtgacatcgytacaacagtcacaccccagtgccctcccttacctactaa 2045
QY 4733 tgaagcaaaacttgyagaltggyaataacttggccaggttaaatgcaggtlaccaatgt 4792
Db 2046 tgaagcaaaacttgyagaltggyaataacttggccaggttaaatgcaggtlaccaatgt 2105
QY 4793 cactatcctcctgtgtggycacatctcagaattttaatgaagaattcttaagacgttctc 4852
Db 2106 cactatcctcctgtgtggycacatctcagaattttaatgaagaattcttaagacgttctc 2165
QY 4853 taaagactatatgtgacatatgtctgaagaagatctggaattgaggyaaaacaacatttcctg 4912
Db 2166 taaagactatatgtgacatatgtctgaagaagatctggaattgaggyaaaacaacatttcctg 2225
QY 4913 atcaggtccagatctgtgttttccaatctccatatgattcttctgatactgtgaaaatga 4972
Db 2226 atcaggtccagatctgtgttttccaatctccatatgattcttctgatactgtgaaaatga 2285
QY 4973 aggaacacacagaggaatcctgaaatgtttgatgtactaatgtacagaatltgccaaggtt 5032
Db 2286 aggaacacacagaggaatcctgaaatgtttgatgtactaatgtacagaatltgccaaggtt 2345
QY 5033 accagacctctccaga--gcgattgacctgtgtgcagaacatgycagccaagcactcag 5089
Db 2346 accagacctctccagatctgcgattgacctgtgtgcagaacatgycagccaagcactcag 2405
QY 5090 aacgaagcaatcatgtctgaagctgcacagtgcttagtccactcagcagcactgttgtctg 5149
Db 2406 aacgaagcaatcatgtctgaagctgcacagtgcttagtccactcagcagcactgttgtctg 2465
QY 5150 aatatttgagcatgtctgaggaacccggaataatcttccgtgtgagatgttaacatttcaga 5209
Db 2466 aatatttgagcatgtctgaggaacccggaataatcttccgtgtgagatgttaacatttcaga 2525
QY 5210 atatttcataatgtttttagaagaatctgcggtctcagatgattgtgtatactccagatg 5269
Db 2526 atatttcataatgtttttagaagaatctgcggtctcagatgattgtgtatactccagatg 2585
QY 5270 aagaaggtatctgctctggaataactttaactgagtcaggaactgtggaattacttgaac 5329
|||||

Db 2586 aagaaggtatctgtctctggaataactttaactgagtcaggactgtggaattacttgaac 2645
QY 5330 aagcagctgtctcctctctatagtctgycgatgtatgaagcaagttaatgaagttacaag 5389
Db 2646 aagcagctgtctcctctctatagtctgycgatgtatgaagcaagttaatgaagttacaag 2705
QY 5390 tactattcctattcatlgaagctaactcgggatgcaagaanaactatccacaattcatgta 5449
Db 2706 tactattcctattcatlgaagctaactcgggatgcaagaanaactatccacaattcatgta 2765
QY 5450 aacttcaagaagcatctcagcaaaattgtttcatccaggtactgctggygcgagtgttg 5509
Db 2766 aacttcaagaagcatctcagcaaaattgtttcatccaggtactgctggygcgagtgttg 2825
QY 5510 gcacctatttcgtgtgtgttttlatggaaccaaagttcgggatttgatgaaacaagaat 5569
Db 2826 gcacctatttcgtgtgtgttttlatggaaccaaagttcgggatttgatgaaacaagaat 2885
QY 5570 ttgtttaaagagcctgcataataaccaaactgcagagatalctcacagattgaggat 5629
Db 2886 ttgtttaaagagcctgcataataaccaaactgcagagatalctcacagattgaggat 2945
QY 5630 tttaacggaagaagatttgagagagatgtgttgaagtaatcaaaagacttcaatcctgtag 5689
Db 2946 tttaacggaagaagatttgagagagatgtgttgaagtaatcaaaagacttcaatcctgtag 3005
QY 5690 acaagtgtaaatttagatccctaaacaaggcataatctcagattacctaattgtgagccatact 5749
Db 3006 acaagtgtaaatttagatccctaaacaaggcataatctcagattacctaattgtgagccatact 3065
QY 5750 ttgacacatatgagatggaaggaacagaatcacacctatttcgacaaaaatlaaatcttcgctc 5809
Db 3066 ttgacacatatgagatggaaggaacagaatcacacctatttcgacaaaaatlaaatcttcgctc 3125
QY 5810 gattcatgtlactgtacaccccttactttagatgycggtgcccattggyggaacttcatgaac 5869
Db 3126 gattcatgtlactgtacaccccttactttagatgycggtgcccattggyggaacttcatgaac 3185
QY 5870 aattcaaaaaggaagacacattctgactacgtctcagtcgcttccctatataatlaaaaaaagg 5929
Db 3186 aattcaaaaaggaagacacattctgactacgtctcagtcgcttccctatataatlaaaaaaagg 3245
QY 5930 tcaatgtcactcataaagaagagatcatcttaacccaattgaagttgctatlgagaca 5989
Db 3246 tcaatgtcactcataaagaagagatcatcttaacccaattgaagttgctatlgagaca 3305
QY 5990 tgcagaaaaaagacacagaggttggcatttggcaacacatcagatcccgcgagaccccaaaa 6049
Db 3306 tgcagaaaaaagacacagaggttggcatttggcaacacatcagatcccgcgagaccccaaaa 3365
QY 6050 tgcctcagatggtactccagggatctgttaggcaccacagtgaaatcaggygccttggaa 6109
Db 3366 tgcctcagatggtactccagggatctgttaggcaccacagtgaaatcaggygccttggaa 3425
QY 6110 ttgcccagggttctctgtctgaaataacctagtgaccccaagctcttcagacatcataata 6169
Db 3426 ttgcccagggttctctgtctgaaataacctagtgaccccaagctcttcagacatcataata 3485
QY 6170 aactcgcactcgtgctttaaagaatttactlaaaagtgltgaagatgccttlaagaaaaata 6229
Db 3486 aactcgcactcgtgctttaaagaatttactlaaaagtgltgaagatgccttlaagaaaaata 3545
QY 6230 agagcttaattggygcggttcaaaaaggtatcaaaaggaattggyggaactatcttcgc 6289
Db 3546 agagcttaattggygcggttcaaaaaggtatcaaaaggaattggyggaactatcttcgc 3605
QY 6290 cttaaagagggccctacagccctagatcacagaagaagtcctcagttatccaagccagtaatt 6349
Db 3606 cttaaagagggccctacagccctagatcacagaagaagtcctcagttatccaagccagtaatt 3665
QY 6350 gcttgctccctgccaacagagattccttcagtgcaatgagcttccgcaaaatygatctcta 6409
Db 3666 gcttgctccctgccaacagagattccttcagtgcaatgagcttccgcaaaatygatctcta 3725
|||||

QY	6410	aactgaaatgcactgttttatttcacctcgcgaagaagccatglatccaacatcgagtgtgaa	6469
Dd	3726	aactgaaatgcactgttttatttcacctcgcgaagaagccatglatccaacatcgagtgtgaa	3785
QY	6470	aagatctatggaaaaccaacaatgynaatgnaatctctgynaattatatcatatggaanaatg	6529
Dd	3786	aagatctatggaaaaccaacaatgynaatgnaatctctgynaattatatcatatggaanaatg	3845
QY	6530	cagtgccaagaanaatacaaatgtagattgttaacgcttgagaatcatatgctatggtt	6589
Dd	3846	cagtgccaagaanaatacaaatgtagattgttlaacgcttgagaatcatatgctatggtt	3905
QY	6590	ctaattgtcgggtaacacaagctgttatctttaagacattttaatgactcaaaaggtacact	6649
Dd	3906	ctaattgtcgggtaacacaagctgttatctttaagacattttaatgactcaaaaggtacact	3965
QY	6650	atacatttacccattatttataccatagctaaaggtlaaaaaatttatccacttaagtctgt	6709
Dd	3966	atacatttacccattatttataccatagctaaaggtlaaaaaatttatccacttaagtctgt	4025
QY	6710	attttttaattatataccaacctttatagatccaattcttgggaaccttttaaatgtagtaat	6769
Dd	4026	attttttaattatataccaacctttatagatccaattcttgggaaccttttaaatgtagtaat	4085
QY	6770	gcttattttaaaggtaactaaaaaataatgtgaatgttttaacctcgtgcgcgccaggcctc	6828
Dd	4086	gcttattttaaaggtaactaaaaaataatgtgaatgttttaacctcgtgcgcgccaggcctc	4144

RESULT	4
AAH41953	
ID	AAH41953 standard; DNA; 4200 BP.

AC AAH41953;

DT 30-AUG-2001 (first entry)

DE CLASP-3 nucleotide fragment.

Human; cadherin-like asymmetry protein; CLASP; CLASP-3; immune response; cell surface molecule; transmembrane protein; immunosuppressive; vaccine; antiinflammatory; antiarthritic; antianaemic; dermatological; uropathic; ophthalmological; antirheumatic; nephrotropic; antithyroid; antidiabetic;; neuroprotective; antiasthmatic; antibacterial; antisense therapy; gene therapy; chromosome 1p31.1; ds.

OS Unidentified.

WO2001.42297-A2.

PD 14-JUN-2001.

PF 13-DEC-2000; 2000WO-US34171.

PR 13-DEC-1999; 99US-0170453.

PR 14-FEB-2000; 2000US-0182296.

PR 11-APR-2000; 2000US-0196460.

PR 11-APR-2000; 2000US-0196528.

PR 13-OCT-2000; 2000US-0240503.

XX 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040

XX

XX

PI Lu P, Garman JD, Candia AF;

DR WPI; 2001-375003/39.

PT Novel isolated cadherin-like asymmetry protein (CLASP) -3, useful for

PT useful in preventing or treating a CRASP-3-mediated disease e.g.
PT autoimmune disease -
XX
PS Disclosure; Fig 9C; 189pp; English.

The present invention describes an isolated polypeptide (I) comprising a nucleotide (nt) sequence (SIa) that has at least 90% identity to the sequence given in AAH41934 and is immunologically cross-reactive with the derived amino acid (aa) sequence (SIb) given in AAB99495 or shares a biological function with native CLASP-3. (I) has immunosuppressive, antiinflammatory, antiarthritic, antianaemic, dermatological, uropathic, ophthalmological, antirheumatic, nephrotropic, antithyroid, antidiabetic, neuroprotective, antiasthmatic and antibacterial activities, and can be used in antisense therapy, vaccine production and gene therapy. CLASP-3 related sequences can be used in preventing or treating a CLASP-3 mediated disease, preferably an autoimmune disease by inhibiting an immune response. The autoimmune disease is caused or exacerbated by increased activity of T helper 1 lymphocytes (TH1). AAH41893 to AAH41953 and AAB99491 to AAB99507 represent sequences which are used in the exemplification of the present invention. CLASP-3 is localised in the chromosome location 1p31.1.

Sequence 4200 BP; 1262 A; 844 C; 907 G; 1187 T; 0 other;

Query Match	60.48;	Score 4123;	DB 22;	Length 4200;
Best Local Similarity	99.98;	Pred. No. 0;		
Matches 4136; Conservative	0;	Mismatches	0;	Indels 3; Gaps 1;

[illegible][illegible]

QY 2813 ttaacactggtgtccaaaagctgccccatgaggatccaaacccagttccaagtgcagaat 2872
|||||
Db 181 ttaacactggtgtccaaaagctgccccatgaggatccaaacccagttccaagtgcagaat 240

QY 2873 caacacagctatgatacgaaattgttaatcgtatgtcttcgcacacagagacgtcaagt 2932
|||||
Db 241 caacacagctatgatacgaaattgttaatcgtatgtcttcgcacacagagacgtcaagt 300
|||||

```
QY      2933   tcttacaacattaacgagcgtttaccactaaagcttttcacgagagctgctt 2992
          |||||
Db       301   tcttacaacattaacgagcgtttaccactaaagcttttcacgagagctgctt 360
```

[illegible][illegible]

Db	481	ctccaagaaagtcgtttctcagaacglttcattgatgacattgcagctctgtcagca	540
xy	5115	ctccaaggaagaacgcgtttcttcaggaacglttcacatgacacatcgcagccctcgcacca	5177

	CGATTCGCTAGTGTACGATTGCAGAGCAGCAAGAATGTTGAGAGCCCA	600
Db		
	CGATTCGCTAGTGTACGATTGCAGAGCAGCAAGAATGTTGAGAGCCCA	720
X7		

Db 601 atacaagccttgcatctcttctcaatgatctgtgtgtctgtatgtgacagaggtttc 660

Db 661 ttagcctataaagtcctgtataaacaggtgtcttccaagcttactcattaccgaatc 720

QY 3353 ccagtgctcgtgtcccttgagcgctggaatttctacgaatcatctgcagtcatgagcact 3412
Db 721 ccagtgctcgtgtcccttgagcgctggaatttctacgaatcatctgcagtcatgagcact 780
QY 3413 atgtacatctaacttacccctgcagcttacttactccacctgcacatctcacctcttg 3472
Db 781 atgtacatctaacttacccctgcagcttacttactccacctgcacatctcacctcttg 840
QY 3473 ttctctgcacaatctcagagttcttgatttctacgaatgttacaagaccaaaagattg 3532
Db 841 ttctctgcacaatctcagagttcttgatttctacgaatgttacaagaccaaaagattg 900
QY 3533 caaatagtttgaaattatccgtgccttccgcgcaaacagcattatttggcaggacttgt 3592
Db 901 caaatagtttgaaattatccgtgccttccgcgcaaacagcattatttggcaggacttgt 960
QY 3593 taacagagctggtctgtcatctttagaaccctgattgctgaagagctgttggattgcataa 3652
Db 961 taacagagctggtctgtcatctttagaaccctgattgctgaagagctgttggattgcataa 1020
QY 3653 aagtcatacaatagtgtacacaatttactctccagtcacgactcagacccgcgtactctg 3712
Db 1021 aagtcatacaatagtgtacacaatttactctccagtcacgactcagacccgcgtactctg 1080
QY 3713 accctcagataaaggctcgagtggtgcacgttgtatctacccctcgtattgtattatcatg 3772
Db 1081 accctcagataaaggctcgagtggtgcacgttgtatctacccctcgtattgtattatcatg 1140
QY 3773 aaactgtaccctcagctgtatgatctttagaagaactcacaaatcaacgaggaagaccaatt 3832
Db 1141 aaactgtaccctcagctgtatgatctttagaagaactcacaaatcaacgaggaagaccaatt 1200
QY 3833 gtagagccactgattgattatgaaagtgagagcggaaglatgataagccagaccgttgcca 3892
Db 1201 gtagagccactgattgattatgaaagtgagagcggaaglatgataagccagaccgttgcca 1260
QY 3893 tggcaatcgcaaggagacatcggtccctcaactaaacaaggcctgagcttccctcctacgt 3952
Db 1261 tggcaatcgcaaggagacatcggtccctcaactaaacaaggcctgagcttccctcctacgt 1320
QY 3953 caacgagtgcgagcgcaacacactaccttctcagcagaatcaagtcgcaagcctttgatct 4012
Db 1321 caacgagtgcgagcgcaacacactaccttctcagcagaatcaagtcgcaagcctttgatct 1380
QY 4013 gtctactttggtgtctcaaaaatgacagtgaaaacagttctacagaagtgtttacagatc 4072
Db 1381 gtctactttggtgtctcaaaaatgacagtgaaaacagttctacagaagtgtttacagatc 1440
QY 4073 tctcagctcttgcaagctaaacggctattagatctgcttatactctcgtgtgtcttgcttg 4132
Db 1441 tctcagctcttgcaagctaaacggctattagatctgcttatactctcgtgtgtcttgcttg 1500
QY 4133 agtataaagggaaaaaagtgcttgaaacgaatgaatagcttgacccttaagaatacaaaag 4192
Db 1501 agtataaagggaaaaaagtgcttgaaacgaatgaatagcttgacccttaagaatacaaaag 1560
QY 4193 acatgagagcaaaagcttgaagaagctatctctgggagcataggtgcagcgcaagaatgg 4252
Db 1561 acatgagagcaaaagcttgaagaagctatctctgggagcataggtgcagcgcaagaatgg 1620
QY 4253 tacggcgaagccgagagacagctcgagagagcccatctggaagtgccttggaaqtcaag 4312
Db 1621 tacggcgaagccgagagacagctcgagagagcccatctggaagtgccttggaaqtcaag 1680
QY 4313 aaaatttgaggtgaggaagaatatagtactcactggtctcaaaaacacagagaagcttgaca 4372
Db 1681 aaaatttgaggtgaggaagaatatagtactcactggtctcaaaaacacagagaagcttgaca 1740
QY 4373 aatcaagagcagagatttgaacacgaaagcactgtattgtgaaacctggtctacagaagcaa 4432
Db 1741 aatcaagagcagagatttgaacacgaaagcactgtattgtgaaacctggtctacagaagcaa 1800
QY 4433 acctaatcatctttagatacatctagagattgtgttcagaccgcttctgttaacggaatcca 4492

Db 1801 acctaatcatctttagatacatctagagattgtgttcagaccgcttctgttaacggaatcca 1860
QY 4493 aagagaagcattctgtgtggaagtgctaaaagtctactacacagcattgacctgtlaaccaa 4552
Db 1861 aagagaagcattctgtgtggaagtgctaaaagtctactacacagcattgacctgtlaaccaa 1920
QY 4553 gtcagtttatctacaacactgttctgtcacacagagagccttggttcaaaagttccctg 4612
Db 1921 gtcagtttatctacaacactgttctgtcacacagagagccttggttcaaaagttccctg 1980
QY 4613 aactctatttgaagaagagacagagcagtggtgtgatttatgtcctcagcgctctccgac 4672
Db 1981 aactctatttgaagaagagacagagcagtggtgtgatttatgtcctcagcgctctccgac 2040
QY 4673 actgtagcagtagcatcgtgtacaatacaggtcaacccccagtgacctcttaactactaa 4732
Db 2041 actgtagcagtagcatcgtgtacaatacaggtcaacccccagtgacctcttaactactaa 2100
QY 4733 tgaaggcaaaactttgagattgggaataacttgcaggggttaaaatgcaggtaccaaagt 4792
Db 2101 tgaaggcaaaactttgagattgggaataacttgcaggggttaaaatgcaggtaccaaagt 2160
QY 4793 cactatccctcctgtgtggacacatctcagaattttaatgaaagattcttaagacgtctc 4852
Db 2161 cactatccctcctgtgtggacacatctcagaattttaatgaaagattcttaagacgtctc 2220
QY 4853 taaagaactatatgacatatgcttgaagaagatctggaattgagggaaacaacatltccctg 4912
Db 2221 taaagaactatatgacatatgcttgaagaagatctggaattgagggaaacaacatltccctg 2280
QY 4913 atcaggtccagatctgttttcaatctccatatgattcttctgatactgtgaaaatga 4972
Db 2281 atcaggtccagatctgttttcaatctccatatgattcttctgatactgtgaaaatga 2340
QY 4973 aggaacaccagagagatccctgaaatgtttgatattgaatctaatgttacagaattgccagggt 5032
Db 2341 aggaacaccagagagatccctgaaatgtttgatattgaatctaatgttacagaattgccagggt 2400
QY 5033 accagacctctccaga---gcgattgacctgtgtgcagaacatgycagcgcaagcactcag 5089
Db 2401 accagacctctccagatctgcgattgacacctgtgtgcagaacatgycagcgcaagcactcag 2460
QY 5090 aacgaaagcaatcatgtctgaaagctgcacagtgctcagtcacactcagcagcactgttgctg 5149
Db 2461 aacgaaagcaatcatgtctgaaagctgcacagtgctcagtcacactcagcagcactgttgctg 2520
QY 5150 aatatttgagcatgtctgagagaccggaatatctctcctgtggagatgtgtaaacatttcaga 5209
Db 2521 aatatttgagcatgtctgagagaccggaatatctctcctgtggagatgtgtaaacatttcaga 2580
QY 5210 atattcatctaagtttttagaagaatctgcggtctcagatgatgtgtatctccagatg 5269
Db 2581 atattcatctaagtttttagaagaatctgcggtctcagatgatgtgtatctccagatg 2640
QY 5270 aagaaggtatctgctctggaaaaatacttactaggtcaggaactgtgtggattactggaac 5329
Db 2641 aagaaggtatctgctctggaaaaatacttactaggtcaggaactgtgtggattactggaac 2700
QY 5330 aagcagctgtctcctctctatgtgctgycatgtatgaagcagyltaatgaagtttacaag 5389
Db 2701 aagcagctgtctcctctctatgtgctgycatgtatgaagcagyltaatgaagtttacaag 2760
QY 5390 tacttatctcatttcatalgaagctaacgcggtgcagaagaactataccacaattcatgta 5449
Db 2761 tacttatctcatttcatalgaagctaacgcggtgcagaagaactataccacaattcatgta 2820
QY 5450 aacttcaagaagcattcagcaaaatgtttcatcagagtagtgcgtgggagcgagatgttg 5509
Db 2821 aacttcaagaagcattcagcaaaatgtttcatcagagtagtgcgtgggagcgagatgttg 2880
QY 5510 gcaacctatttctgtgtgttttlatggaaccaagttcggggatttggatgaacaagaat 5569

Db 2881 gcacctatttcgtgtgttttattatggaaccaagtlcggggatttgatgaacaagaat 2940
QY 5570 ttgttcaaggagcctgcataaaccaaacttgcagagatatccacagattggaggat 5629
Db 2941 ttgttcaaggagcctgcataaaccaaacttgcagagatatccacagattggaggat 3000
QY 5630 ttacggagaaagatttggagagatgtgttgaagtaatcaaaagactctaaccctag 5689
Db 3001 ttacggagaaagatttggagagatgtgttgaagtaatcaaaagactctaaccctag 3060
QY 5690 acaagtgaatattagatccctaacaggcatatatccagatracctatgttgagccact 5749
Db 3061 acaagtgaatattagatccctaacaggcatatatccagatracctatgttgagccact 3120
QY 5750 ttgacacatatgagatgaaggacgaatcacctatttcgacaaaaattacaatctcgtc 5809
Db 3121 ttgacacatatgagatgaaggacgaatcacctatttcgacaaaaattacaatctcgtc 3180
QY 5810 gatcatgtactgtacaccccttactttagatggccgtgcccatygggaacttcagtaac 5869
Db 3181 gatcatgtactgtacaccccttactttagatggccgtgcccatygggaacttcagtaac 3240
QY 5870 aattcaaaaggaagaccattctgactacgtctcatgtgccttccttataattaacaagg 5929
Db 3241 aattcaaaaggaagaccattctgactacgtctcatgtgccttccttataattaacaagg 3300
QY 5930 tcaatgtcactcataaagaagatcatcttaacaccaatitgaagtgtctatttggaca 5989
Db 3301 tcaatgtcactcataaagaagatcatcttaacaccaatitgaagtgtctatttggaca 3360
QY 5990 tgcagaaaaagacacagaggttgcatttgcacaacatcagatcccgagaccccaaaa 6049
Db 3361 tgcagaaaaagacacagaggttgcatttgcacaacatcagatcccgagaccccaaaa 3420
QY 6050 tgcctcagaatgttactccagggatctgttagggcaccacagtgaaatcaggggccttgyaag 6109
Db 3421 tgcctcagaatgttactccagggatctgttagggcaccacagtgaaatcaggggccttgyaag 3480
QY 6110 ttgcccaggttttctgtctgaatacctagtgaccccaagctcttcagacatcataata 6169
Db 3481 ttgcccaggttttctgtctgaatacctagtgaccccaagctcttcagacatcataata 3540
QY 6170 aactgcgactctgctttaaagatttactaaagtgltgaagatgccttaagaaaaata 6229
Db 3541 aactgcgactctgctttaaagatttactaaagtgltgaagatgccttaagaaaaata 3600
QY 6230 agagcttaattggccggttcacaaaggagtatcaaaaggaattggggaactatcttcgc 6289
Db 3601 agagcttaattggccggttcacaaaggagtatcaaaaggaattggggaactatcttcgc 3660
QY 6290 cttaagaggccctacagccctagatcacagaaagtcctcagttatccaagccagtatt 6349
Db 3661 cttaagaggccctacagccctagatcacagaaagtcctcagttatccaagccagtatt 3720
QY 6350 gctgtccctgcacacagagatlccttcagtcgaatgagctttcgcaaatggtctcta 6409
Db 3721 gctgtccctgcacacagagatlccttcagtcgaatgagctttcgcaaatggtctcta 3780
QY 6410 aactgaatgcactgtttatcatctgcaaaagagccatgtattcaacatcgagtgtgaa 6469
Db 3781 aactgaatgcactgtttatcatctgcaaaagagccatgtattcaacatcgagtgtgaa 3840
QY 6470 aagatctattggaaccaacatggaatgtgaaattctggaattatattcatitgaagaatg 6529
Db 3841 aagatctattggaaccaacatggaatgtgaaattctggaattatattcatitgaagaatg 3900
QY 6530 cagtgccaagaataatatacaatgtagattgttaacgcttgagaatcatatggtctgttt 6589
Db 3901 cagtgccaagaataatatacaatgtagattgttaacgcttgagaatcatatggtctgttt 3960
QY 6590 ctaatgttcgggtacaagctgttatacttttaagacatlttaatgtactcaaaaggtacact 6649
Db 3961 ctaatgttcgggtacaagctgttatacttttaagacatlttaatgtactcaaaaggtacact 4020

QY 6650 atacattaccattattataccatagctaaaggttaaaaaatttacttaacttaagttcgt 6709
Db 4021 atacattaccattattataccatagctaaaggttaaaaaatttacttaacttaagttcgt 4080
QY 6710 atttttaattataccacatttatagattcatttggaccacattttaaatgtagtaaat 6769
Db 4081 atttttaattataccacatttatagattcatttggaccacattttaaatgtagtaaat 4140
QY 6770 gcttatttaaggtactaaaaaatatgtgaatgtttacctcgtgcgcgacaggcctc 6828
Db 4141 gcttatttaaggtactaaaaaatatgtgaatgtttacctcgtgcgcgacaggcctc 4199

RESULT 5
AAH41952
ID AAH41952 standard; DNA; 3173 BP.
XX
AC AAH41952;
XX
DT 30-AUG-2001 (first entry)
XX
DE CLASP-3 nucleotide fragment.
XX
KW Human; cadherin-like asymmetry protein; CLASP; CLASP-3; immune response;
KW cell surface molecule; transmembrane protein; immunosuppressive; vaccine;
KW antiinflammatory; antiarthritic; antianaemic; dermatological; uropathic;
KW ophthalmological; antirheumatic; nephrotropic; antithyroid; antidiabetic;
KW neuroprotective; antiasthmatic; antibacterial; antisense therapy;
KW gene therapy; chromosome 1p31.1; ds.
XX
OS Unidentified.
XX
PN WO200142297-A2.
XX
PD 14-JUN-2001.
XX
PF 13-DEC-2000; 2000WO-US34171.
XX
PR 13-DEC-1999; 99US-0170453.
PR 14-JAN-2000; 2000US-0176195.
PR 14-FEB-2000; 2000US-0182296.
PR 11-APR-2000; 2000US-0196267.
PR 11-APR-2000; 2000US-0196460.
PR 11-APR-2000; 2000US-0196527.
PR 11-APR-2000; 2000US-0196528.
PR 11-APR-2000; 2000US-0547276.
PR 13-OCT-2000; 2000US-0240503.
PR 13-OCT-2000; 2000US-0240508.
XX
PA (ARBO-) ARBOR VITA CORP.
XX
PI Lu P, Garman JD, Candia AF;
XX
DR WPI; 2001-375003/39.
XX
PT Novel isolated cadherin-like asymmetry protein (CLASP) -3, useful for
PT useful in preventing or treating a CLASP-3-mediated disease e.g.
PT autoimmune disease -
XX
PS Disclosure; Fig 9A; 189pp; English.
XX

The present invention describes an isolated polypeptide (I) comprising a nucleotide (nt) sequence (SIa) that has at least 90% identity to the sequence given in AAH41934 and is immunologically cross-reactive with CC the derived amino acid (aa) sequence (SIb) given in AAB99495 or shares a CC biological function with native CLASP-3. (I) has immunosuppressive, CC antiinflammatory, antiarthritic, antianaemic, dermatological, uropathic, CC ophthalmological, antirheumatic, nephrotropic, antithyroid, antidiabetic, CC neuroprotective, antiasthmatic and antibacterial activities, and can be CC used in antisense therapy, vaccine production and gene therapy. CLASP-3 CC related sequences can be used in preventing or treating a CLASP-3 CC mediated disease, preferably an autoimmune disease by inhibiting an

CC immune response. The autoimmune disease is caused or exacerbated by
CC increased activity of T helper 1 lymphocytes (TH1). AAH41893 to AAH41953
CC and AAB99491 to AAB99507 represent sequences which are used in the
CC exemplification of the present invention. CLASP-3 is localised in the
CC chromosome location 1p31.1.

Sequence 3173 BP; 981 A; 624 C; 688 G; 880 T; 0 other;

Query Match	44.68;	Score 3044.8;	DB 22;	Length 3173;
Best Local Similarity	97.18;	pred. No. 0;		
Matches 3155; Conservative	0;	Mismatches 12;	Indels 83;	Gaps 22;

OY	3582	aggacttgtgttaacagagctgctgtcatl	tttagaccctgatgctgaag	actgtttgg	3641
Db	3	aggacttgtgtl	ttcagagctgctgtcatl	tttagaccctgatgctgaag	actgtttgg 62

Qy	3642	atgcataagaagtcatacaatattactctccagtcacgactcagacc	3701
Db	63	attgcataagaagtcatacaatattactctccagtcacgactcagacc	122

OY	3702	gcggtactctgacccctcagataaaagctcgagtggccatgttgtatctaaccttgattgg	3761
Db	123	gcggtactctgacccctcagataaaagctcgagtggccatgttgtatctaaccttgattgg	182

OY		3762	tattatcatggaactgttacccctcagctgttatgatltttacagaactcaccaatcaaacyy	3821
Dd		183	tattatcatggaacctgtaccctcagctgttatgatltttacagaactcaccaatcaaacyy	242

QY 3822 aagacccaatttgtatagccactgtatgtattatgaaagtgaagcggaagtatgataagcca 3881
|||||
Db 243 aagacccaatttgtatagccactgtatgtattatgaaagtgaagcggaagtatgataagcca 302

QY 3882 gaccgttgcacatgcacatcggtcccttaactaacaaggcctgcaattt 3941
|||||
Db 303 gaccgttgcacatgcacatcggtcccttaactaacaaggcctgcaattt 362

QY 3942 cctcctcagtcacagagtggcagcgcaacacactactccttttcagcagaatcaagtcgaag 4001
|||||
db 363 cctcctcagtcacagagtggcagcgcaacacactactccttttcagcagaatcaagtcgaag 422

```
QY    4002 cctttgatctgtcacttgggttcctcaaaaatgcagatgaacagtctctacagaagtg 4061
      |||||
Db    423 cctttgatctgtcacttgggttcctcaaaaatgcagatgaacagtctctacagaagtg 482
```

```

2Y 4062 gttacagatctctcagctcttgcaagtaaacgcgctattagatctgtcttattctctgtg~4121
      |||||
Db 483 gttacagatctctcagctcttgcaagtaaacgcgctattagatctgtcttattctctgtg 542

```

```

QY      4122 gtcttgcttggagataaaagggaaaaagtgtttgacgaatgaatatagcttgacctttaa 4181
          |||||
Db      543   gtcttgcttggagtataaagggaaaaagtgtttgacgaatgaatatagcttgacctttaa 602

```

QY 4182 gaatcaaaagacatgagagcaaaagcttgaagaagcattcttggagcatagtgtccag 4241
|||||
Db 603 gaatcaaaagactagagagcaaaagcttgaagaagcattcttggagcatagtgtccag 662

4242 gcaagaatatgttaccgycgaagccgagacagctcagagaaagcccatctggaagtgcct 4301
 |||||
 Db 663 gcaagaatatgttaccgycgaagccgagacagctcagagaaagcccatctggaagtgcct 722

QY 4302 tgggaagtcagaagaattttaggttgaggaagatagtactactgtgctcaaacacaga 4361
|||||
Db 723 tgggaagtcagaagaattttaggttgaggaagatagtactactgtgctcaaacacaga 782

Dy 4302 gaagcttgacaaatcgaagcagagattgacaacgcgaactgatgtgaaacctggc 4421
|||||
Db 783 gaagcttgacaaatcaagaagcagagattgaacacgcgaagccactgtattgtgaaacctggc 842

Dy 4422 tacagaagcaaacctcatcatttagatacacatitagagtgtgttcagaccggtttcgt 4481
|||||
Db 843 tacagaagcaaaagttaacattttagatatacattagagtgtgttcagaccggtttcgt 902
|||||

QY	4482	aacggaatccaaagagagcatcttgtgtgagtgctaaagctctactacacagcatgyc	4541
Db	903	aacggaatccaaagagagcatcttgtgtgagtgctaaagctctactacacagcatgyc	962
QY	4542	ctgttaaccaaagtgtcagttatctacacacactgttttgcctacacagagagccttgtttc	4601
Db	963	ctgttaaccaaagtgtcagttatctacacacactgttttgcctacacagagagccttgtttc	1022

QY	4602	aaagtttcctgaactcttatttgaagaagagacagacagcgtgctgatttatgcctcag	4661
Db	1023	aaagtttcctgaactcttatttgaagaagagacagacagcagcgtgctgatttatgcctcag	1082

QY	4662	gctctccgacactgttagcagtagcatcgtlacaatacgcgtcacaccccaagtcctccct	4721
Db	1083	gctctccgacactgttagcagtagcatcgtlacaatacgcgtcacaccccaagtcctccct	1142

[illegible]

QY 4782 ggtaccaatgtcactatcctccttggtyggcacatctcagaatttaatgaagaattcct 4841
 |||||
 Db 1203 ggtaccaatgtcactatcctccttggtyggcacatctcagaatttaatgaagaattcct 1262

QY 4842 aaagcgttctctaagaactatattgacataatgctgaagaagaatctggaattgagggaac 4901
|||||
Db 1263 aaagcgttctctaagaactatattgacataatgctgaagaagaatctggaattgagggaac 1322

```

QY 4902 aacatttcctgatacaggtccaggaatctggttttcaatctccataatcttcttgataac 4961
      |||||
Db 1323 aacatttcctgatacaggtccaggaatctggttttcaatctccataatcttcttgataac 1382

```

QY 4962 ttgtgaaatgaaggaacacccaggagatccttgaatatgtttgatctatctatgttacagaat 5021
|||||
Db 1383 ttgtgaaatgaaggaacacccaggagatccttgaatatgtttgatctatctatgttacagaat 1442

QY	5022	tgccaaggttaccagacctctccaga---gcgattgacctggttcagaacatgycagg	5078
Db	1443	tgccaaggttacccagacctctcagatctgcgatgtgacctggttcagaacatgycagg	1502

QY	5079	caagcactcagaacgaagcaatcatgtctgaagctgcacagtgcttagtccactcagcagc	5138
Db	1503	caagcactcagaacgaagcaatcatgtctgaagctgcacagtgcttagtccactcagcagc	1562

QY 5139 actgtgtcgtgaataatttgagcatgctgtaggaccggaataatcttcctgtggaatgtgt 5198
 |||||
 Db 1563 actgtgtcgtgaataatttgagcatgctgtaggaccggaataatcttcctgtgccatgtgt 1622
 |||||

QY 5199 aacattcagataatttcacccaatgttttagaagaatctgcggtctcagatgatgtgt 5258
 |||||
 Db 1623 aacattcagaataattcatctaagtlttagaagaatctgcggtctcagatgatgtgt 1682

Db	1683	atctccagatgaagaagtatctgctctgaaaaataacttacttgtagtcaggactgtggg	1742
QY	5259	atctccagatgaagaagtatctgctctgaaaaataacttacttgtagtcaggactgtggg	5318

Db 1743 attactggaacaagcagctgtcttctctctatgctgctgcatgtatgaagcagttaatga 1802
 |||||
 Uy 5319 attactggaacaagcagctgtcttctctctatgctgctgcatgtatgaagcagttaatga 5378

53/9	agttacaaga	tactatctat	ctatcatgaa	gctaactc	ggaacata	ctccac	5438
1803	agttacaaga	tactatctat	ctatcatgaa	gctaactc	ggaacata	ctccac	1862

y 3439 aatcattgtaaaccttcaagaagcattcagcaaaattgtcttcagagtaactggtcgga 5498
 |||
 Db 1863 aatcattgtaaaccttcaagaagcattcagcaaaattgtcttcagagtaactggtcgga 1922

[illegible]

Db 1983 tgaacaagaatttgtttaacaaggagcctgcataaaccaaacttgcaagagatatctcacag 2042

QY 5619 attggaaggattttaacggaagaagaatttggagaaggatgtgttgaagtaatcaaaagactc 5678

Db 2043 attggaaggattttaacggaagaagaatttggagaaggatgtgttgaagtaatcaaaagactc 2102

QY 5679 taatcctgtagaacaagtgtlaaattagatccctaacaaggcatatatctcagattacatagt 5738

Db 2103 taatcctgtagaacaagtgtlaaattagatccctaacaaggcatatatctcagattacatagt 2162

QY 5739 ggaagccatacttgcacacatatgagatggaagacagaatcacctatttcgacaaaatta 5798

Db 2163 ggaagccatacttgcacacatatgagatggaagacagaatcacctatttcgacaaaatta 2222

QY 5799 caatctcgtcgcattcatgttactgttacaccccttactttagatggccgtgcccattggga 5858

Db 2223 caatctcgtcgcattcatgttactgttacaccccttactttagatggccgtgcccattggga 2282

QY 5859 acttcatgaaacaattcaaaaaggaaagaccattctgtactacgtctcattgccccttccctatat 5918

Db 2283 acttcatgaaacaattcaaaaaggaaagaccattctgtactacgtctcattgccccttccctatat 2342

QY 5919 taaaacaagggtcgaatgtcactcataaagaagagatcatcttaacccaattggaagtgc 5978

Db 2343 taaaacaagggtcgaatgtcactcataaagaagagatcatcttaacccaattggaagtgc 2402

QY 5979 tatlgagacatgcagaaaaagacacagagttggtcatttgcaacacatcagatccgcgc 6038

Db 2403 tatlgagacatgcagaaaaagacacagagttggtcatttgcaacacatcagatccgcgc 2462

QY 6039 agaccccaaaatgtctcagatgtgtactccagggatctgttaggcaccacagtgaaatcaggg 6098

Db 2463 agaccccaaaatgtctcagatgtgtactccagggatctgttaggcaccacagtgaaatcaggg 2522

QY 6099 gcctttggaagtggcccaagtttttctgtctgtaaatatacctagtgaacccaagctcttcag 6158

Db 2523 gcctttggaagtggcccaagtttttctgtctgtaaatatacctagtgaacccaagctcttcag 2582

QY 6159 acatcataataaactgcgactctgtcctttaaagatttactaaaagtggtgaagatgcctt 6218

Db 2583 acatcataataaactgcgactctgtcctttaaagatttactaaaagtggtgaagatgcctt 2642

QY 6219 aagaaaaaataaaggcttaattgggcccgttcaaaaaggagtaaccaagggaattggggaa 6278

Db 2643 aagaaaaaataaaggcttaattgggcccgttcaaaaaggagtaaccaagggaattggggaa 2702

QY 6279 actatctcgcccttaaaaggagccctacagagccctagatcacagaagaagtcctcagttatcc 6338

Db 2703 actatctcgcccttaaaaggagccctacagagccctagatcacagaagaagtcctcagttatcc 2762

QY 6339 aagccagtatgtctgtcccccgtgcacagagattccttcagtcgaatgagctttcgcaaa 6398

Db 2763 aagccagtatgtctgtcccccgtgcacagagattccttcagtcgaatgagctttcgcaaa 2822

QY 6399 atggaatccttaaacatgcaatgcaactgtttatcatctgcaaaagagccatgtattcaaca 6458

Db 2823 atggaatccttaaacatgcaatgcaactgtttatcatctgcaaaagagccatgtattcaaca 2882

QY 6459 tcgaagtgtgaaaaagatctatttggaaaaccaatggaatctctggaataattatattc 6518

Db 2883 tcgaagtgtgaaaaagatctatttggaaaaccaatggaatctctggaataattatattc 2942

QY 6519 attgaagaatgcagtgcccaagaaatalcaaatgtlagaatgtltaacgcttgagaatcat 6578

Db 2943 attgaagaatgcagtgcccaagaaatalcaaatgtlagaatgtltaacgcttgagaatcat 2959

QY 6579 ggcatagtttctaatgttgcggttaacaagcgttatctttaaagacattttaatgactc 6638

Db 2960 -----ctttaaagacattttaatgactc 2982

QY 6639 aaaggtacactatacatattacattattatataccatagctaaggttaaaaaattattcac 6698

Db 2983 aaaggtacactatagatattacattattatataccatagctlaaggtlaaaaaattattcac 3042

QY 6699 tttaaagtcgtatttttaaatatatatcacattatagattcatcttgagccatttta 6758

Db 3043 tttaaagtcgtatttttaaatatatatcacattatagattcatcttgagccatttta 3102

QY 6759 aatgtagtaatgcttatttttaaaggtactaaaaaataatgtgaatgttaccctcgtgcgcg 6818

Db 3103 aatgtagtaatgcttatttttaaaggtactaaaaaataatgtgaatgttaccctcgtgcgcg 3162

QY 6819 ccagggcctc 6828

Db 3163 ccagggcctc 3172

RESULT 6

AAH43851

ID AAH43851 standard; cDNA; 6372 BP.

XX

AC AAH43851;

XX

DT 04-SEP-2001 (first entry)

XX

DE Human CLASP-7 encoding cDNA sequence SEQ ID NO:1.

XX

KW Human; CLASP-7; cadherin-like asymmetry protein; immune response; neuroprotective; antidiabetic; immunosuppressive; antirheumatic; antiarthritic; hypotensive; anti-HIV; cytostatic; immunostimulant; antianaemic; antinflammatory; ophthalmological; nephrotrophic; antithyroid; antiasthmatic; antiallergic; antibacterial; gene therapy; chromosome 19q13.2; autoimmune disease; multiple sclerosis; toxemia; juvenile diabetes; rheumatoid arthritis; pruritic urticarial papule; hypertension; Rh incompatibility; ss.

KW

KW

XX

OS Homo sapiens.

XX

XX

FH Key Location/Qualifiers

FT CDS 13..6156

FT /*tag= a

FT /product= "CLASP-7"

XX

PN WO200142295-A2.

XX

PD 14-JUN-2001.

XX

PE 13-DEC-2000; 2000WO-US34152.

XX

PR 13-DEC-1999; 99US-0170453.

PR 14-JAN-2000; 2000US-0176195.

PR 14-FEB-2000; 2000US-0182296.

PR 11-APR-2000; 2000US-0196267.

PR 11-APR-2000; 2000US-0196460.

PR 11-APR-2000; 2000US-0196527.

PR 11-APR-2000; 2000US-0196528.

PR 11-APR-2000; 2000US-0547276.

PR 13-OCT-2000; 2000US-0240503.

PR 13-OCT-2000; 2000US-0240508.

PR 13-OCT-2000; 2000US-0240539.

PR 13-OCT-2000; 2000US-0240543.

XX

PA (ARBO-) ARBOR VITA CORP.

PI Lu P, Garman JD, Candia AF;

XX

DR WPI; 2001-381641/40.

DR P-PSDB; AAB99541.

XX

PT Novel cadherin-like asymmetry protein-7 and polynucleotides encoding the polypeptide, useful for treating autoimmune disease, hypersensitivity, preventing transplant rejection by modulating immune response

PT -

XX

PS Claim 1; Fig 5; 151pp; English.

XX The present invention describes a human cadherin-like asymmetry protein (CLASP), designated CLASP-7. The CLASP-7 protein (I) and its encoding nucleotide sequence (II) have activities including: neuroprotective; anti-diabetic; immunosuppressive; antirheumatic; antiarthritic; anti-HIV; hypotensive; cytostatic; immunostimulant; antianaemic; antiinflammatory; ophthalmological; nephroretrophic; antithyroid; antiasthmatic; antiallergic; and antibacterial. (II) and CLASP-7 antibodies (III) are useful for detecting the CLASP-7 polypeptide. (II) is useful for producing (I) by recombinant methods. (I) or its fragment are useful for inhibiting an immune response in a cell such as T cell or B cell. A pharmaceutical composition (C), comprising (I) or (II), can be useful for treating CLASP-7-mediated disease such as an autoimmune disease caused or exacerbated by increased activity of T helper cells. Autoimmune diseases which can be treated using (C) include multiple sclerosis, juvenile diabetes and rheumatoid arthritis. (I) is useful for treating toxemia or pregnancy induced hypertension, pruritic urticarial papules and Rh incompatibility. (I) is also useful as a diagnostic reagent for immune and other disorders, since diseases characterised by overproduction or depletion of lymphocytes in blood or other organs may be detected by monitoring the level of (I) or its mRNA. CLASP-7 has been mapped to the chromosomal location 19q13.2. The present sequence encodes the human CLASP-7 protein as given in the present invention.

Sequence 6372 BP; 1273 A; 2050 C; 1849 G; 1200 T; 0 other;

Query Match	31.2%;	Score 2132.8;	DB 22;	Length 6372;
Best Local Similarity	60.4%;	Pred. No. 0;		
Matches 3863;	Conservative	0;	Mismatches 2262;	Indels 273;
				Gaps 10;

Qy	26	ccgagcgccgcgccttcgccagaagatcagcagaacggtgycagccgaagttagaagc	85
Db	23	ccgagcgccgcgccttcgcgcacaagatcaaaagacggtgycgcgacagagtgcggaagc	82

QY 86 agatctccgacatatagltgltctccccaactgtcaaaaacctaatattgttgca 14
 || ||| | || ||||| || ||||| || |||
 Db 83 aggtgtcccggaacgcagtgctctcccccacattcagcagcgctgcagcagctc----- 13

```

QY 146 atatacccatcaccacagtgccccttacggaagcagtagatccagltgatttgaag 205
      |  |||||  |||||  |||||  |||||  |||||  |||||
Db 138 -----ctggggtcccaatgactgaagtgtgcagccccctggaacttgaag 184

```

OY 206 attaccataactcaatcccttgctgtgattctgggccttaacggatttgtaat 265
|| || | | || | ||| | ||| || | |||
Db 185 atgtacttctgagccggcaccagaagtgtagcccgggcccttcagggaacctgtagaat 244

QY 266 ttcttcacagatgatatgtgaagtgtttataagtcctcggagactgcagaactcttgttcag 325
| | | |||| | | | | | | | | | |
Db 245 tcccaagtgaagacttggagctgtctctgcagccccgggaatgcccgaccacgagagccc 304

QY 326 ctgtacctgaagaaagtgaatgatccacatgtagactgtataagaagttatacag 385
| | | | | | | | | | | | | | |
Db 305 g gatcccccaagatgaaaaacttgatgccaggctgaaggcgccggtgagatgtatatg 364

QY 386 aagactggygcattgtcatcagaataatcataaattggyaacaggaatttaalccaata 445
| | | | | | | | | | | | | | | | | | | | | |
Db 365 aggactggytcaattgtccacagaagtatcagtaacctgagtgcagcatatacagcccggtca 424

QY 446 cattagataaacacagaagaagcaccaaaagtgttcaccaaaacaagtttctgaatctgaty 505
 | ||| |||| | | | | | | | | | | |
Db 425 ccacagacacacacgcgagcgacacagaaggccctccccgcccaagtccttgagcacgagt 484

QY 506 aaqctccagatgycacaacagctaccagatgatcaagaatgacottaaaagaacgttcaatgt 565
 || ||| | | ||| | |||
 Db 485 ctctcgagacgaagagctccgcgcctfagagactcgaatgattccgcgcgtgtgctcgcgct 544

QY 566 caatagatgatacccccaaggggtagctggcctctgtagtatactttgacttgaaaaattcac 625
 | | | | | | | | | | | | | | | | | |
 Db 545 ccccggaagacacccctcgaaagcaggggtgcctctagcatcttcgacctgaggaactg 604

QY 626 ttccgtatgcttctgtcttcccaattacttgatcgaaactccaaatgaagaatagaccgtc 685
||||| ||||| ||| ||| ||| ||| ||||| |||||

Db 605 cagctgactcatgtgtgccctctctgtctagagcggcgccccaagaatatgtgaccgc 664

QY	686	agaatgatgaccaaaggaatccaacccgtcacaaagaacttttgccttgcatac	745
Db	665	gcaatgaaaccccttcgacgcgcagcaccgcccgcgccttgcacccttaccgcgcac	724

OY 746 cagatgagggaagaaccaataagaacgcgttagtgttcctgatataccaagaacatttg 805
 | | | | | | | | | | | | | | | | | | | | | |
Db 725 ctgacgagatgtgaagccgttgtgaacgctgttagccgccagagcacaaccgcgcgagcaactttg 784

QY	806	gtcaaaagcactctgttaa	atbctta	caactcaagtttga	aatgaaatgaa	atgaaaccat	tt	865
Db	785	gacaaagcatctgtgtca	agtgctgtcgctcaag	ttcgagattgaa	atgaaatgagcc	atct		844

QY 866 ttgcaagtttggttatatattgcaagaaagaatttcagaaaactttatt 925
||| | ||||| ||||| ||| ||| ||||| |||
Db 845 ttgggattctgctctgtatgatgtcgggagaaaaagaagattctcgagaacttctact 904

QY 926 ttgaccttaattctgagcagatgaagggttglttaagtcacatgtaccacctgtgcga 985
| | | | | | | | | | | | | | | |
Db 905 tcgactggaactcgactccacgaaggggctgtctcggtccaatgatgccaccaacctgcga 964

QY 986 ttactaccctgycgaagatcagaatttttctatactaattacccctccaagatgttttc 1045
| | ||||| | | | | | | | | | | | |
Db 965 tctcacacctggccgcctctgcacattctctgtgacctaacccctcaactgacatcttc 1024

QY 1046 tttaataaagctagaanaagtcctacagcaaggaacatctgtgcagaaccat 1105
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1025 tggatcaagcttgagaaggtcttcagcaagggaacatcagtgatgtgtgagcct 1084

QY 1106 atttgatttccaagaagcagatgccaaccaagaataaagaataactygagaaactgaaaga 1165

Db 1085 acatggtgtgtaagaagatgtaacacagccaagaacacaaagagaagctagagaagctgcgcc 1144

DY 1166 gteaagcagatcagtctttgccaaagacttgggaatatcgcatgaccttgcttgacgtg
||| |||||| | | | ||| ||| ||| ||
Db 1145 tggcgccgcgacgtcttcaccgccctgtggccgctaaccgatgccttcgcctlgacgg 1204

QY 1226 caatccatttaatgaatatctgttagcagtcgtcgtggagttctgaaagagattctacagaag 1285
 1205 ccgtgcacttggccaacatcgtgagcagcgtcgtggcagctgacccggactctgactcgg 1264
 Db

Qy	1286	tagaatacgtactgagagaacaaaaaggtctctgtgcagatagagggaattctagtattg	1345
Db	1265	ag-----gycgagcgccgcccagccttgacagaccgccc-----	1298

QY 1346 ttgacagacgacacttgaaagacaacaagtggagatgatctgttaacttgacgagct 1405

Db 1299 -ccgtcggggccccagacccggcgagtagtgggagcagcagctgcagcttctctgct 1357

QY 1406 ttccgaccagctactctcacagtgacaatttttttaagcaggaagagaccgccttaagt 1465

Db 1358 ttccgtccagccagctactctgtcacaaacttctttaagcagagagctgagcgactcagt 1417

QY 1466 atgaagatctctacaatttcctgtcgatagagaagccatcttctgtcttaacygcgac 1525

Db 1418 acgagacctctcaagttctcgtcgatagagcgcgccgtctccctgctgcgcgac 1477

Qy 1526 taagacctattacagctcagctcaagatagacattctccgcaccttgaaaatcccccat 1585

Db 1478 taagtcctgtgactgccagctcaagatcgacattctcgcgtcctgaaaatcccccat 1537

QY 1586 atgcctaactccgagctgtcttcaagtgaagctttaccctcgacagtagatttagacctc 1645
 ||||| - || ||||||||| | |||| | |||| | | |||
 Db 1538 tctgcctctccctgagctgtcttcatatcaagccctaccgagaccaccagggcgcccca 1597

QY 1646 ccagagaatcttagagttcccgcaaggagatgtttatgttccaacactacttacagaa 1705
 ||| ||| | ||||| ||||| | ||| ||| |||||
 Db 1598 ccaaggagatcttgagttcccgcccgcaagtcatagtcccccataccagctacagga 1657

QY 1706 atctctctacatataccctcagagtccttaattttgccaaatgcgaagttctgtctagaa 1765
 Db 1658 acctgctgtacgtgtaccgcgacagcctcaacttcagcagcgcgcgaaggtcttcgtgcga 1717

QY	1766	ataaacagtgtaaaagtcacggttlatglatgagagatccaagcaatgcccgttaa	1825
Db	1718	acctgtgtgcaggtgcaglacatgacagcgagacccacgacgctctgcccgtca	1777
QY	1826	tcttgttaaatcctagctgtltcagaaatttcacaaagaaagcctatacagccgtatatac	1885
Db	1778	tcttggcaagtcacagctgcagtgaaatttaccgcgagggccttcacacccggtgtctacc	1837
QY	1886	ataacaggtctcctgatttctcatgaaaatcaaggttaaagcttcctgtctacttaactg	1945
Db	1838	ataacaggtcccccgagttctacgagaggttcaagcgcgcattctccagcctgcgtgacag	1897
QY	1946	accatcatcactgtccttttactcttttatcatglttagttgtcaacaaaaacaatactc	2005
Db	1898	agaaccatcacctgtgtttcaccttctaccatgttcagctgcagccccgcgcgcaactg	1957
QY	2006	ctcttgaaacaccagtttgatatacatgataccaatgcttcagaatgagcgttgaaga	2065
Db	1958	ccctggaagacacccgctgggcttacttgatcccaactgtctgacgacggcgcttgaaga	2017
QY	2066	ctggccaagtttctgtccagctcctatgtgaaaaaaccaaccaaggtatctgtactgt	2125
Db	2018	ccggccctctgtctcccaagtgctgtgacacagccgcgccagctattccggtctca	2077
QY	2126	ctcctgaggtctcctctaactgycatgaaatggtgataaatacacaaggtgttttaatg	2185
Db	2078	cacccgatgtgcgcttccgycgatgcgtggtgagcgttcacaagggcgtgttcaagt	2137
QY	2186	ttgaagttgtgtgttcgtctatccatatcacacaagatccttatcttgacaattttg	2245
Db	2138	tggaagctacacagccggtgtcctctgtgcacccccagagacccctactgacaaattctca	2197
QY	2246	ctctgtcaatgctctgagatgaacacacctgttccagctccgaattggggacatgcgaatca	2305
Db	2198	ccctgtgtcacgtccctggaaggagagagccttcccaattccgctcaagagacactgtgcga	2257
QY	2306	tggaataataacttagaanaatgaattgaagagcagfatattcagcactgaattcatcccaagc	2365
Db	2258	gcgagggcaacgctggagcagagagctgcgggccagctcttgacagcactgcgcctggccagcc	2317
QY	2366	tggaaccagtggtccgaattcttctcatcttctgtcgtagataaaactgatacttttagttata	2425
Db	2318	ccgaacccctgtgtgccttctcccaaccaagctgtcgtgacaaagctcgtgcgtgtcatca	2377
QY	2426	gaacctcgtcatgtctgycgccaataagttaaacctagtcgaagcatctttgaagccaatgg	2485
Db	2378	ggcccccgatcatcagtgycgaagattgtgaacctgggcccgttgaagcaatgg	2437
QY	2486	catcaattataaatcgactcacaaaacttgaagaagaaatcatgaccagcatgycagaa	2545
Db	2438	cccatgtagtcagcctgtgttccacgcgagcctggaagagcccaagatgccccgcgtcact	2497
QY	2546	acagccctctgtcatcatatatatcatatgtttccgcctaccaaatacttaacctaat	2605
Db	2498	gccacacagctgctgcctacgtccactacgacgccttcgccttctgtgcactgagccca---	2554
QY	2606	catcatcacccaggtccctggggggttggagagatcagtgcatatatgccaacatggtcagat	2665
Db	2555	-----gcctcccgagatggggccccctccagtgacagtgagcgtgccaacactggccccgtg	2608
QY	2666	ctgcggtgagacctgcaagccttaatttaaatcgttcttcgaagccttagtaataagcaatc	2725
Db	2609	gctctgtgcgcccgcaagcctctacactgycgcgttccaagaagcatcaagcagcagcaacc	2668
QY	2726	cagatatatactggagactcccaacgctcaccaagatgatgaagttcgatcaatcatcgggagta	2785
Db	2669	ctgacctgcgcgttggccccctgctctgtgtgatacgaaggttcccgcatcctgcccag--	2726
QY	2786	agggtttagatcgtcccaattcctggttacaacatggtgttccaaaagctgcccccatggg	2845
Db	2727	-----	2726

[illegible]

[illegible]

Db	4745	tcaticgacctcatgttacaagaatttgcccggggcttaccaggctcacccggaccttcgctgtg	4804
QY	5057	cctgtttgcagaacaatggcagcgcaagcactcagaaacgaaacatcatgttgaagctgcac	5116
Db	4805	cctggttgcagaacaatggtccgggaagcagcgagagctgtggcaaccacgcgcgagccgcc	4864
QY	5117	agtgctagtcaccactcagcagcactgtgtgctgtaatatlttgacgatgcttggagccgga	5176
Db	4865	agtgcatgtgtgcacgcgcgcgccttcgtgtgtgtgtacctgcgcctgtctcgaggaccagc	4924
QY	5177	aatatcttcctgttgggagtgtgaacatttcagatatltcatctaattttaagaagaat	5236
Db	4925	gccacctgcccgttggctgcgttcttccttcagagacatctcatccaacgttgcTlagagagt	4984
QY	5237	ctgcggtctcagatgagtgtgtatctccagatgaaagaagtatctgctctgaaataact	5296
Db	4985	ccgcacatctccgacgacacatctgttcgcgcgcgagggaggtctctgctccgggaagcact	5044
QY	5297	ttactgagtcaggaacttgttgggattactgtgaacaaagcagctgtcttcctctctatgctg	5356
Db	5045	tcactgagctgggctgtgtaggttgtgtggaacagcgagccggtacttaccatgggcg	5104
QY	5357	gcatgtatgaagcaggttaatgaagtttacaagaacttattcctatctcatgaagctaac	5416
Db	5105	ggctctacgagcggtgtagtgaagtgctacaagaaccttaccatcctgtgaagcccaacc	5164
QY	5417	gggagtcaaaagaacctaccacaattcatgtgtlaacctcaagaagcaltcagcaaaatlg	5476
Db	5165	gtgactacaagaagctgycgcgcggtgcacgycaaactcgcagagcgcttcaccaagatca	5224
QY	5477	ttcatcagagtactgctgctgggagcggaatgttttggcacctatttctgtgtgtttatg	5536
Db	5225	tgcaaccagagttccggctctgggagcgcggtgttcgggacgtatltccgcgttggcttcaag	5284
QY	5537	gaaccaagtctcggggaatttggatgaaacaagaatttgtttacaaggaagcctgcataacca	5596
Db	5285	gcgcgccacttcggtgtgacctgtgatgagcaggaagtttgtgtacaaggaagccatcgatacga	5344
QY	5597	aacttgcagagatatctcacaagattggagggaattttaacgggagaaagatttggagagagt	5656
Db	5345	agctgycagagatctcacaccgcgttggaggagtctctacacgcgagagatttggcgcgcgcg	5404
QY	5657	tggttgaaagtaatcaaaagactctaactcctgttagacaagtgtaaatlagatcctaacaag	5716
Db	5405	tcgttgagattatcaaaagactcttaccctgttgacaagttccaagcttgactcacaaaag	5464
QY	5717	catatatcagattacctaattgttggagccatacttggacacatatgagatgaagyacagaa	5776
Db	5465	ccctacatccagatactacgtatgttggaaaccgttacttgtatactaagagctcaagaccggg	5524
QY	5777	tcacctatttcgacaaaatatfacaactcttcgtcgaattcatgtactgtlacacccttactt	5836
Db	5525	tgaacttactttgaccgcgaactatgtggcttcgcacattccctgttctgcacgcgcgttcaagc	5584
QY	5837	tagatggccgtgcccattgggggaacttcatgacaacaattcaaaaggaagaccattctgacta	5896
Db	5585	cggatgycgcgcgacacacggggagctgcccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc	5644
QY	5897	cgctcatgcttcccttatataataaacaaggytcaatgtcactcataaagaagagatca	5956
Db	5645	ccgacacagccttccctacatcaagaactcgcattccgtgtgtgccaccgagagagacgg	5704
QY	5957	tcctaacaccaattgaagttgcttaattgagacatgacagaaaagacacaggaattggcat	6016
Db	5705	tgtctagcgcagtgagagtgtygcacatcgagagacatgcagaagaagacacggagctgacct	5764
QY	6017	ttgcacaacatcaggaatcccgacgagaccccaaatgtcttcagatgttactccaggaatcg	6076
Db	5765	ttgcacaccgagcaggaacccacagatgtctaagatgtctacagagtggtcttcagggtctg	5824
QY	6077	taggcacacacagtgaaatcaggggccttttggaagttgcccaggttttctgtctgaaatac	6136
Db	5825	tagggccacacggtgaaccacaggttccctctggaagtgtygccacggtgtttttagcagagatcc	5884

QY	6137	ctagtgcacccaaagctcttcagacatcataataaactgcgactctgctttaaagattta	6196
		1 1111 11111111 1111 11 111 111111 11 11	
Db	5885	cggaagaccccaagctcttcgycatcacacaacaattgcgctctgtcttcaagyaacct	5944
QY	6197	ctaaagtggtgaagatgccttaagaaaaataagacttaattgggccggttcaaaag	6256
		11 1 1111 1111 1 1111111 1 11111111 11 1111	
Db	5945	gcaagaatatgtgagatgcgctgcgaaaaataagccctgatttggccgcgacccgaag	6004
QY	6257	agatcaaaaggaattg--ggaaactatcttcgccttaagagagccctacagccctaga	6314
		1111 11 11 11 11111 11111 1111 11 11111 1	
Db	6005	agtaaccacgctgtagcgcggagcgcaactactgcgcgctgcggagagctctgcagccctgc	6064
QY	6315	tcacagaaagtccttcagttatccaagccagtatgtgtccctgcacagagattcc	6374
		1 11 1 11 1 1 111 11 1 111111 11 111 111	
Db	6065	ttaccacagcgcttgcgcccaactgtatgycaccccaccc---accgcgctcaggaactcc	6120
QY	6375	ttcagtcgaatgaagcttttcgcaaaatgatgactctaaac	6412
		11 1 11 11 11 11 11 11111 11 1111 11	
Db	6121	ttgaacagagcaagttccgaaaagcgacagcctctgaac	6158

RESULT	7	
ID	AAS07373	
AC	AAS07373;	
DT	26-SEP-2001	(first entry)
DE	Human cDNA encoding CLASP-5.	
XX	Human; CLASP-5; cadherin-like asymmetry protein; immune gateway; immunogen; antibody; autoimmune disease; rheumatoid arthritis; multiple sclerosis; leukaemia; insulin dependent diabetes mellitus; acquired immunodeficiency syndrome; AIDS; ss.	
OS	Homo sapiens.	
XX		
FH	Key	Location/Qualifiers
FT	CDS	112..6048
FT		/*tag= a
FT		/product= "CLASP-5"
FT	primer_bind	1498..1519
FT		/*tag= b
FT		/note= "PCR primer HC5gs1"
FT	primer_bind	3087..3113
FT		/*tag= c
FT		/note= "Antisense oligonucleotide AAS07384"
FT	primer_bind	3087..3113
FT		/*tag= d
FT		/note= "Antisense oligonucleotide AAS07384"
FT	primer_bind	3232..3259
FT		/*tag= e
FT		/note= "Antisense oligonucleotide AAS07384"
FT	primer_bind	complement (3642..3660)
FT		/*tag= f
FT		/note= "PCR primer HC5As10b"
FT	primer_bind	4728..4756
FT		/*tag= g
FT		/note= "Antisense oligonucleotide AAS07384"
PN	WO200142296-A2.	
XX		
PD	14-JUN-2001.	
XX		
PF	13-DEC-2000;	2000WO-US34163.
XX		
PR	13-DEC-1999;	99US-0170453.
PR	14-JAN-2000;	2000US-0176195.
PR	14-FEB-2000;	2000US-0182296.
PR	11-APR-2000;	2000US-0196267.
PR	11-APR-2000;	2000US-0196460.

PR 11-APR-2000; 2000US-0196527.
PR 11-APR-2000; 2000US-0196528.
PR 11-APR-2000; 2000US-0547276.
PR 13-OCT-2000; 2000US-0240503.
PR 13-OCT-2000; 2000US-0240508.
PR 13-OCT-2000; 2000US-0240543.
PR 13-OCT-2000; 2000US-0240539.
XX
PA (ARBO-) ARBOR VITA CORP.
XX
PI Lu P, Garman JD, Candia AF;
XX
DR WPI; 2001-367865/38.
DR P-PSDB; AAU04024.
XX
XX CLASP-5 polynucleotides, proteins and antibodies are used to prevent or
PT treat a CLASP-5 mediated disease, such as an autoimmune disease e.g.
PT Rheumatoid arthritis -
XX
PS Claim 1; Fig 6; 188pp; English.

CC The sequence encodes a cadherin-like asymmetry protein, CLASP-5, which
CC is a transmembrane protein of the immune system involved in the formation
CC of the immune gateway. CLASP-5, polynucleotides encoding it and an
CC anti-CLASP-5 antibody are used to prevent or treat a CLASP-5 mediated
CC disease, such as an autoimmune disease caused or exacerbated by
CC increased activity of TH1 cells. These diseases may include Addison's
CC disease, rheumatoid arthritis, dermatitis, Multiple Sclerosis, Systemic
CC Lupus Erythematosus and Autoimmune Thyroiditis, inflammatory conditions
CC (e.g ischaemia-reperfusion) and responses, Leukaemia, acquired
CC immunodeficiency syndrome, AIDS, syphilis, typhoid, Burkitt's lymphoma,
CC Grave's disease and insulin dependent diabetes mellitis. CLASP-5 is used
CC for inhibiting an immune response in a cell (T cell or B cell) by
CC interfering with the expression of a CLASP-5 gene in the cell, the
CC ability of a CLASP-5 protein to bind to another cell or the ability of a
CC CLASP-5 protein to bind to another protein. The antibody binding CLASP-5
CC is used to inhibit an immune response is a subject. The polynucleotide is
CC used to detect CLASP-5 expression in cells and for diagnosis of
CC diseases and disorders associated with aberrant expression of CLASP-5.
XX
SQ Sequence 7215 BP; 2005 A; 1734 C; 1721 G; 1755 T; 0 other;

Query Match	25.6%	Score 1749;	DB 22;	Length 7215;
Best Local Similarity	58.28;	Pred. No. 0;		
Matches 3613; Conservative	0;	Mismatches 2360;	Indels 238;	Gaps 21;

[illegible]

Db 425 ccgctgctcagggcccccgcacctaaacgtgtgtgcgacgtgtctgggaaagcc 484

QY 590 gctgggacctgtagtatctttgacttgaaaaattcacttcctgatgtcttgcctccaatt 649

Db 485 ccgtcaactgctgtgactttgacctccgcagccttgcaagccttgcaagcggtagaanaacc 544

QY 650 tactgtatcgaaactccaatatgaaagaatatagaccgtcagaatgtgaccaaaggaatacaa 709

Db 545 tccctgcagcaagtgtgcaggagactttgagaagcagaacgagagggcccgaggaacca 604

QY 710 accgtcacaaagaacttttgctttgcatccatcacccagatgaggaagaaaccaatagaac 769

Db 605 acaggcagggccgagctcttgccctttaccatcagtgagcagagagagatgctgtgga 664

QY 770 ggcttagtgtctcgtatatacccaagaacatttggtcaaaagactcttgrtaaatgct 829

Db 665 tacgtccagtaaccagaatgtcccaaggaaacacctgggcaacagaatatattggtcaagtgc 724

QY 830 tatcaactcaagtttgaatatgaaattgaaccatttttgcaagtttgctttatatgatg 889

Db 725 tgaccttgaaagtcgagattgaaaattgagccctgtttgccaagcattgcccctaacgatg 784

QY 890 tcaaggaaaaagaaaaagattcagaaaaacttlatittgaccttaattctgagcagatga 949

Db 785 ttaagaaagaagaaaaagatctcagaaaatttcactgtgacctgcaactctgaccagttca 844

QY 950 aagggtgttaccgtccacatgtacacactgtgccaactactaccctggcaagcagacgcaa 1009

Db 845 aaggattctgcgagctcacagccttcaagtggccgcacatcaagtcagggcagatcgcag 904

QY 1010 ttttctatcactatcctcccaagatgttttctgttaataaagctagaanaagtc 1069

Db 905 tcttcacgtacactaccgctccacagacatctaccctgtagtcagaatttgaaanaagtc 964

QY 1070 tacagcaaggagacatltggagagtggtgcagaaccatatatgatttccaagaagcagatg 1129

Db 965 tgcagcagggagatatltggagactgtgcagagccctacacggttatcaagaagaagtgatg 1024

QY 1130 ccaccaagaataaagaaaaactggagaaactgaagagtlcaagcagatcagttttgccaaa 1189

Db 1025 gtggaagaagataaagaaaagattgaaaactaaactccaagctgaaatccttctgcagc 1084

QY 1190 gacttgggaaatatcgcacgtcccttttgcttggactgcaatccatttaatgaatatgtta 1249

Db 1085 gtttgggaaatalccggaatgcccttgccttgcctggcaccctaagcttatacaagcttct-- 1142

QY 1250 gcaagtctgggagtttggaaagagattctacagaagtagaanaatcagtlactggagacgaa 1309

Db 1143 -caatgtctccacccttggagggaggtlaactgttga-----ctctgtgttggga 1195

QY 1310 aagggtctgtgcagagaggaattctagtattgttgcagacagatcacttgaagga 1369

Db 1196 gaagcccaagtgggtgaacggagagacatltggccaatctagaaggcttc--tgaagag 1252

QY 1370 caacaagtggagatgtatgcttgytaacttgcagcagcttccagccagctactctcacagtga 1429

Db 1253 ccctctccttggaggaanaatggggttggatccaacttcaaaaactcctcactctgagcgta 1312

QY 1430 caaatlttttaagcaggaaggaacccgcttaagtgtatgaagatactctacaatctcttg 1489

Db 1313 gcagcttlttcaagcaggaagagatcgcccttagcgatgaagaacttatacaagttttag 1372

QY 1490 ctgatatgagaagggcatctctgtcttaagcgcgactaagacataltacagctcagctca 1549

Db 1373 ctgactacaanaagatcatcatccttaccagagcagatcaagtcaaltccaggttgtaaa 1432

QY 1550 agatagacatttctccgcacgtgaanaatccccaatlatttgccaactccgagctgcttc 1609

Db 1433 gactggagatttctacagctccagagatcatcaattgtctgtcagctcctgaaatgtcgc 1492

QY 1610 aagtgaagctttaccctgacagtagagttagacctacccagagaanaatcttagagttcccg 1669

Db 1493 ccgtgaacccttctcctgaaaaccgcgacacgcccgcacaaagagattttggaatttccaa 1552

QY 1670 caaggatgttaltgttccaaacactacttacagaanaatctctctacatataccctcaga 1729

Db 1553 cagcagaagatatatgtccctcacactgtgtacagaanaacctctctatgttctaccacaga 1612

QY 1730 gtcttaatttgcocatcgtcaaggttctgtctagaanaatataacagtgaaagtcagttta 1789

Db 1613 ggtgaactttgttaacaaactagcatcagcccggacaatataaagaatccagttta 1672

QY 1790 tgtatggaagatcccaagcaatggccatgcccgtatcttgyttaatctagctgttcag 1849

Db 1673 tgttggagaagatgttagcaatgcatgcccgtcatcttgyaaaatccagcgggctg 1732

QY 1850 aatttcaaggaagcctatacacgccgtagtatatcatacaaggtctcctgatttcatg 1909

Db 1733 aattctgcaggaagtgtaacacagctgttatacatacaataaagtlctcctgactttatg 1792

QY 1910 aagaatcaaggttaagcttccctgctaactttaaactgaccatcaatcagttttaact 1969

Db 1793 aagaagtgaanaattaaagctcccgctaagctcacagtaaatcacacctcctgttcaact 1852

QY 1970 ttatcatgttagtgttcaacaaaacaaataactcctcttgaacacacagttgatarata 2029

Db 1853 tctaccatatcagctgttcagcagaagaagagccctccgttgaanaactctcctgggatat 1912

QY 2030 catgataccaatgtctcagaatggagcgttgaagactgcccagtttgcctgcccagct 2089

Db 1913 catgctgcceaatctcttaaatgacgtcttcaactgtaatcctactgtctccagttg 1972

QY 2090 cattgaaaaaaccaaccacagcgttatctgttactgtctcctag-----gttctc 2140

Db 1973 ccttggaaaaaatgtgccaccacactactccatgtcatctgtctvgagaagtlcccatlaaga 2032

QY 2141 tacttggcataaatagggtagataatcacaaagtglttlttaagtltgaagttgtgtg 2200

Db 2033 atcctcccatlaagtlggcgtgaagacataaaggagtaltttaatatltgaagtgcagctg 2092

QY 2201 ttctgtctatcatacacaagaatccttatacttgcacaaattttgtcctgtgtcaatgtc 2260

Db 2093 ttctctgtlacaacaccagacacacacactggagaagttcttcaacctctgcacatcc 2152

QY 2261 tggatgaacacctg---ttccagctccgaattggggacatgcyaatcatggaanaact 2317

Db 2153 tggagagccagtggaaccttcccatccgctgtgtgtgatcagaaanaatcagcyagatggcg 2212

QY 2318 tagaaatgaatgtgaagcagatatttcaagcactgaaatltcatccagctggaaccagtgg 2377

Db 2213 tggagcatgagctggaagctcagcatcatctgtcctgtaactcctccgcttggagccgctcg 2272

QY 2378 tccgattcttcatccttctgtcgtagataaactgatactttagttattagacctcctgtca 2437

Db 2273 tgccttctcgtcacacctgtgtgtgcagaagctcttccagctgtccgtgcagcccatgttca 2332

QY 2438 ttgttggccaataagttaaactagtlcaagcatcttltgaagccaatgcatcaatlataa 2497

Db 2333 tcgctggccagacagccaactctcccaagtttgccttcgagtcggtggtggccatgcgca 2392

QY 2498 atcgaactcacaaaaacttggaaagaaatcatgaccagcagcatgycagaanaagccttctg 2557

Db 2393 acagctgcacacaacagcaagacctgagccaaggaaccagcatggggaactgctgctgg 2452

QY 2558 catcatatatcatatagtlttccgccttacaanaatcacttaccctaat---tcatcatcac 2614

Db 2453 ctctcctagtgcaactagcttctccgccttgcagagaggtgcaaaaggagatgtgcccaagtca 2512

QY 2615 caggtccttgggggttgggagatcagtlcagttatgtccacaatgtgtagatctgcygtga 2674

Db 2513 gcgctcccaactgcccctccttagaccctcgagagctaacacaagatgtggccgacatcagctg 2572

QY 2675 gaactgcaagccttaatttaaatcgttctcgaagccttagtaalatgaatccagatatat 2734

Db 2573 ctgctgtgagttcaaaagctgtgcagggcccggtgtgatgagcagcaagtaaccagacctg 2632

QY	2735	ctgaggactcccaagctcaaccaagatgataagttcgaatcaatcaatcgcggagtaagggttaag	2794
Db	2633	cggggaacaacactccgcagcagacgaggaagtgaagaacatcatgtcttcaagaatcgccg	2692
QY	2795	atcgctccaattccctggtgttaaacactgtgtgtccaanaagctgcccatgaggatccaacc	2854
Db	2693	atcgcaactgcagc-----	2706
QY	2855	ccagtcacaagtgcagaatcaacacacaggcgtatgatacgaaagtgttaatcgtatgtcttcgc	2914
Db	2707	-----cgaaagtcttaactattgtcctctgcga	2731
QY	2915	acacagagacgctcaagttctcttaacaacattaacggygagcgtttaccactaaaaagcctt	2974
Db	2732	gtagtgatgtctccaagtt-----cacctgcagccccaagggccaagccaanaaagcatt	2785
QY	2975	ttcacgaggaagctgagcttcttcagtggtgtgtttgcagtgycagcglttcggaatcagctt	3034
Db	2786	ttcatgaggaagctgtcccttcagaatgtgtgtgcagaccgy-----	2825
QY	3035	tgcaacaagcctgttcttttttgaaatgaatggtlaaagagcagtgtgcaccatttaact	3094
Db	2826	-----aatggtgaaaagcatgycgccagcagctaacata	2857
QY	3095	ttaatgataaacttgaggtctccaaggaanaagtcgtttccagaacggtttcatgtatgaca	3154
Db	2858	acatgacacaaacgagcagagtttcggaagactcgtttcttcgaccggtttcatgtatgaca	2917
QY	3155	ttgcagctctgtcagcaacgattgtcgtatgatatagtttcaagatt-----tcaga	3205
Db	2918	taactactattgttaatgtgtgtcacctcggaaatttgcagcccttttagtaaaaccacaga	2977
QY	3206	aggcacacagaatatggttgagagagactcaatacagaacgcttgcatcttcttccaatgactgt	3265
Db	2978	aggaataatgaaacagcgcggaanaagatgaaacatcagccttgcgttcttctgtatgaccttc	3037
QY	3266	tgctcgttatgacagaggaattgtttttagcccttaaaagtcctgctataaacaggtgt	3325
Db	3038	tctccctcatgatacggyggttgtgtttaaccctcagaacattattgcagccaagctgt	3097
QY	3326	attcaaaagcttactcattaccgaatcccaagtgcttcgtgtccttgagcgctggaatttc	3385
Db	3098	cagccaagctcagtaaccttccaacgctca-----tttcatgagcgtagaagttcc	3148
QY	3386	taagaaatcatctgcagtcacatgagcaactatgttacaataacttaaccctgcagcttaacta	3445
Db	3149	tgagaatcctctgttagcatgagcatattacctcaatctgaacct-----tttttatga	3202
QY	3446	ctccacctgcatactccaatcacctctctgttctcttcgcacaacatctcagaagttctgatttt	3505
Db	3203	atgctgatactgctccaacatctcctctgtctccttccatacttccccaagactcaagctcct	3262
QY	3506	ctacgaatgtacagaaccaaaaagatttgcaaatatgtttgaattatccgtgtccctttccgcc	3565
Db	3263	gctccagcttccagagcaccagaagaatcgccagcagcagtttcgacttccgagtaaccgcc	3322
QY	3566	aacagcattatttgcgcaggaactgtgtgttaacagagctgctgtcatatttagaccctgatg	3625
Db	3323	agcagcaacttctctcaaccgggctcctcttccacagaactgctgtcgtccctgtgatgcggaag	3382
QY	3626	ctgaaggaactgtttcggattgcatagaagaagtcacaatatgtgtacacaatttaactctcoa	3685
Db	3383	gggaaggaatcagcaaaagtacaaaagaagctgttcagltgcaattcacagacgtgttaagtt	3442
QY	3686	gtcacgactcagaccgcggtactctgacccctcagataaaagctcgagtgycatgttgt	3745
Db	3443	ctcacgacctggaaccaacgctgtgttcaaacccagaggtgaaggtcaaaatcgccgcccttt	3502
QY	3746	atctaacctctgattcgttatatcatgtgaaactgttaacctcagctgttatgtatttacagaa	3805
Db	3503	acctaacctttagttggcatcatatttggatgtccttgcacacagcctcgtgtcactttacaagttg	3562
QY	3806	ctcaacaatcaacgaggaagaccaaatttgttataagccaatgatgatattatgaaggtgagagcg	3865

[illegible]

Db 4613 cagccatgcagatgactcctttcccaaccaggtggaagaaacttctctgtaatctgaata 4672
QY 4946 tgattcttctgatactgtgaaaatgaagaaacaccagagatccctgaatatgtgattg 5005
Db 4673 gcatcttatatgacacagagaaatgagggatttccaggaagatccctgatatgt 4732
QY 5006 atctaattgacagaattgccaaggtttaccagacctctccaga---gcgattgacctgt 5062
Db 4733 atctcatgtacagaattgccaagagtaccagggcatctcctgatctgcgctgacctggc 4792
QY 5063 tgcagaaactgycagggcaagcactcagaacggaatcatgtctgaaagctgcacagtgtc 5122
Db 4793 tccagaaacatgycagagaaacacaccagaagaaagtgtcacacggaggctgccatgtgcc 4852
QY 5123 tagtccactcagcagcactgttgcgtgaatatattgagcatgtctggaagaccggaata 5182
Db 4853 tgtgtcacgcgcgctgcgcttaagtgtgctgaatcctgagcatgtctggaagaccacagc 4912
QY 5183 ttccctgtgggatgtgttaacatttcagaatatctcatctaatgttttagaagaatctgc 5242
Db 4913 tgcgcgtggtgcagtggtcagcttccagaatatcttctccaatgtctggaagagtgct 4972
QY 5243 tctcagatgtagtgtgtatctccagatggaagaaagtatctgtctctggaataacttact 5302
Db 4973 tctctgaggaacaccctgtcactgcagcagagatggtgtgcgcagggccaagtactcacc 5032
QY 5303 agtcaggaactgttgggatatcactggaacaagcagctgtctcctctctatgtgtgcgcat 5362
Db 5033 agagtggcctgttaggacctctctggaagcagggccggagctctctcagcacgggaagct 5092
QY 5363 atgaagcagtttaatgaagttatacaaaagtactattccttattcataagcctaaccggat 5422
Db 5093 atgagacagtttaatgaggtctacaagctgtgtcatcccatcctcagaagcgcatacgaa 5152
QY 5423 caaagaactatcccaactcatgtgtaaacctccaagaaagcattccagcaaaatgtgtcac 5482
Db 5153 tccggaagctgacactcatcaccagcaagctgcagagagccttcgacagcatcgtttaaca 5212
QY 5483 agagtactggtcgtggagcggatgttggcacctatttcgtgtgtgtttttatggaacca 5542
Db 5213 aggatcat-----aagagaatgtttggaacctactccgagttgtgttcttggatcca 5266
QY 5543 agttcgggatttggatgaaacaagaattgtttacaaggaagcctgcaataaccaacttg 5602
Db 5267 aatttgggatttggatgaaacaggagtttgtctacaagaagccttgcaattaccaagcttc 5326
QY 5603 cagagatatctcacagattggaaggattttaccgagagaagaatttggagaaggatgtgttg 5662
Db 5327 ctgagatctcacatagactagaggcaattttatgtgtcaatgttttgtgtgcagaattgt 5386
QY 5663 aagtaatcaaaagactctaactcgttagacaagtgtaaattagatccctaacaaggcata 5722
Db 5387 aagtgaattaaagactccaactcgttggacaacaaacgaagttgatccctaacaagccctaca 5446
QY 5723 ttcagattacctatgttggagccatacttggacacatalgagatgaaggacagaataccct 5782
Db 5447 tacagatacaactttgtgagccctacttgtatgaglatgagatgaagaacaggltcacat 5506
QY 5783 atttcgacaaaaattacaatcttcgtcgtatcagtactgttacaccccttactttagatg 5842
Db 5507 acttgagaagaatttcaacctccggagggttcagtgtacacaccccggttcaaccctgtagg 5566
QY 5843 gccgtgcccatggyggaacttcaatcaaatcaaaaggaagaccattctgactaagtcctc 5902
Db 5567 ggcggcctcgggggagagctgtcagtgaagcagtaacagaaggaaacacagltccctgacacat 5626
QY 5903 atgccttcccttatataataaaagggttcaatgttcaactcataaaagaagagatcatcctta 5962
Db 5627 acgccttccctatacatcaagaccaggtatcagcgttcataccagaaggaaggttgttttga 5686
QY 5963 caccaatgaagttgtctatatgagacatgycagaaaaagacacaggaagttgycatttgcaa 6022
Db 5687 caccgatgaagttgcatatgaagaacatgaagaagaagaccctgcagttagcagtttgcca 5746

QY 6023 cacatcagatcccgacagaccccaaaatgtctcagatgttactccagggatctgtaggca 6082
Db 5747 ttaaccaggaagccgcttgatgcaaaagatgtctcagatgtgtctgcaaggctctgtggag 5806
QY 6083 ccacagtgaaatcagggccttltggaagtlgtcccaagttttctgtctgaataactagt 6142
Db 5807 ctactgtaaatcagggaccactggaagtagcccaagtttlttggctgtgaattcctgtctg 5866
QY 6143 acccaaaagctcttcagacatcataataaactgcgactcgtctttaaagatttactaaaa 6202
Db 5867 atccaaaactctatcgcacatcacacaagtltagtgttatgtctttaaagattcatlga 5926
QY 6203 ggtgtgaagatgccttgaagaaaaataagagctttaatgtggccggttcacaaaggatlc 6262
Db 5927 gatgtgtgaagctgttagagaaaaacaagcgtctcatcacygcagaccagaggaatalc 5986
QY 6263 aaagggaaattggyga--aactatcttcgcctttaaagagggccctacagccctlagatc--ac 6318
Db 5987 agcaggaactcaaaagaactataacaagctaaagagaaacctcagggcaatgtatcgagc 6046
QY 6319 agaaaagtcacctcagttatccaaagccagttat 6349
Db 6047 ggaaaattccagaactgtlacaagccaatat 6077

RESULT 8
AAS07382
ID AAS07382 standard; cDNA; 4027 BP.
XX
AC AAS07382;
XX
DT 26-SEP-2001 (first entry)
XX
DE Human DNA associated with CLASP-5 #2.
XX
KW Human; CLASP-5; cadherin-like asymmetry protein; immune gateway;
KW immunogen; antibody; autoimmune disease; rheumatoid arthritis;
KW multiple sclerosis; leukaemia; insulin dependent diabetes mellitis;
KW acquired immunodeficiency syndrome; AIDS; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..2970
FT /*tag= a
FT /product= "protein as displayed in AAU04026"
FT /partial
FT /note= "No start codon"

PN WO200142296-A2.
XX
PD 14-JUN-2001.
XX
PF 13-DEC-2000; 2000WO-US34163.
XX
PR 13-DEC-1999; 99US-0170453.
PR 14-JAN-2000; 2000US-0176195.
PR 14-FEB-2000; 2000US-0182296.
PR 11-APR-2000; 2000US-0196267.
PR 11-APR-2000; 2000US-0196460.
PR 11-APR-2000; 2000US-0196527.
PR 11-APR-2000; 2000US-0196528.
PR 11-APR-2000; 2000US-0547276.
PR 13-OCT-2000; 2000US-0240503.
PR 13-OCT-2000; 2000US-0240508.
PR 13-OCT-2000; 2000US-0240543.
PR 13-OCT-2000; 2000US-0240539.
XX
PA (ARBO-) ARBOR VITA CORP.
XX
PI Lu P, Garman JD, Candia AF;
XX

QY	5057	cctggttcgagaacatgycagggcaagcactcagaaacggaacatcatgtctgaagctgcac	5116
Db	1598	cctgctccagaacatgycagaaacacacacaaagaagtgtctacacagggctgcac	1657
QY	5117	agtgctagtcacactcagcagcactgtgtctgaatatlttagcatgctggaagccgga	5176
Db	1658	tgtagcgtgtgcacgcgcgtcgctagtgcttgaglatctgacatgctggaagaccaca	1717
QY	5177	aatatcttcctgtggagatgttaacattccagaatatctcatctaagtlttaagaagaat	5236
Db	1718	gctaacctgcccgtgycagatgttcagcttccagaatatlttccaatgtgtcggaggaat	1777
QY	5237	ctgcggtcctcagatgtytglatctccagatgaagaaglatctgctctggaataact	5296
Db	1778	ctgtgtctctgagacacccctgtcacctgacggagatyggtgtgycgagggcagtlact	1837
QY	5297	ttactgagtcaggaactgtggatctactggaacaagcagctgtcttcctctctatgctg	5356
Db	1838	tcaccgagagtggcctgtagggcctcctgagcagggccggagctcttcagccgggag	1897
QY	5357	gcattgatgaagcagtttaatgaagtttacaaagtlactatttcctatttcatagaagcta	5416
Db	1898	gcttatatgagacagtttaatgaggtctacaaagctgtgtcatcccatcctagaagcgcat	1957
QY	5417	gggagtcaagaacatctacccaatttcattgtaacttcaagaagcattcagccaataatg	5476
Db	1958	gagaattccggaagctgacactcactcacagcaagctgcagagagccttcgacagcatcg	2017
QY	5477	ttcatcagagtagctgctgctgagcgagtggtttgcacctatttcgtgtgtgtttatg	5536
Db	2018	ttaacaagatcat-----aagaaatgtttggaacctactccgagttgtttcttg	2071
QY	5537	gaaccaagttcgggatttgatgaacaagaatttgtttacaaggaagcctgcataacca	5596
Db	2072	gatccaatttggggatttgatgaacaggaatttctctacaagaagcctgcataacca	2131
QY	5597	aacttcgagagatatctcaccagatttgagggaatttaccggagaaagaattggaagagatg	5656
Db	2132	agcttcctgagatctcacatagactagagcatlttatgtgtcaatgtlttgtgcagaat	2191
QY	5657	tggttgaagtaatcaaaagactctaactcctgttagacaagtgtaaatatagatcctaacaag	5716
Db	2192	ttgtggaagtgttaagaactccactcctgttgacaacaaacgaattgtgatcctaacaag	2251
QY	5717	cataattcagattactctaattgtgagccataactttgacacatatgagatgaagacagaa	5776
Db	2252	cctacatacagatcactlttgtgagccctacttltgatlgaatgagatgaagaacaggg	2311
QY	5777	tcacctattcgacaaaaattacaactctctgcatcatgtactgtacacaccttaact	5836
Db	2312	tcacataacttgaagaagaattcaacctccggaggttcatgttacaccaccccgttcaccc	2371
QY	5837	tagatggccgtgcccattggyggaacttcatagaacaattcaaaagaagaccattctgacta	5896
Db	2372	tggaagggcgcgcctcggygagagctgcatgagcagtlacagaagaacacacagctccctga	2431
QY	5897	cgtctcatgccccttcctatatattaacaagaaggtcaatgtcactcataaagaagagatca	5956
Db	2432	ctatgcacgccttccccctacatcaagaccaggaatcagcgltcatccagaagggaggtttg	2491
QY	5957	tcttaacacccaattgaattgtctatttgagacatgcagaaaaagacacaggaattggcat	6016
Db	2492	ttttgacacccgattgaagttgcatltgaagacatgaagaagaagaccctgcagtttagag	2551
QY	6017	ttgcaaacacatcagatcccggcagaccccaaatgtcttcagatggtactccagggatctg	6076
Db	2552	ttgccaattaaccaggaagccgctgtatgcaaaagatgcttcagatggtgtgcgaagcgtcg	2611
QY	6077	taggcaccacagtgatcaaggggcccttggaaagtgtgccaggttttctgtctgaataac	6136
Db	2612	tgggagctactgttaaatcaagggacacacttggaagttagcccaagtgttttggctgaattc	2671
QY	6137	ctagtgaacccaagctcttcagacatcataataaactgcgactctgtcttaagaatttta	6196

	Dn	2672	ctgctgatccaaactctatcgcacatacacacaagttggagtattgtcttaaggcaattcca	2731
	Qy	6197	ctaaaagtgltgaagatgcccttaagaaaaaataagagcttaattggccggttcaaagg	6256
	Dn	2732	tcatgagatgttgtgaagctgttagaagaaaaaacaaagcgltcatcaacygcagaccagaagg	2791
	Qy	6257	agatcaaagygaattgggga--aactatcttcgaccttaagagagccctacagccctaga	6314
	Dn	2792	aatatcagcaggaactcaaaaagaactataacaagctaagaaggaacctcagaccaatga	2851
	Qy	6315	tc--acagaaagtccttcagttatcccaagccagtatt	6349
	Dn	2852	tcgagcggaataattccagaactgttacaaagccaatatt	2888
RESULT		9		
ID	AAH43850			
	AAH43850 standard; cDNA; 2148 BP.			
AC	AAH43850;			
DT	04-SEP-2001	(first entry)		
DE	Preliminary human CLASP-7 encoding cDNA sequence.			
KW	Human; CLASP-7; cadherin-like asymmetry protein; immune response; neuroprotective; antidiabetic; immunosuppressive; anti-rheumatic; antiarthritic; hypotensive; anti-HIV; cytostatic; immunostimulant; antianaemic; anti-inflammatory; ophthalmological; nephrotrophic; antithyroid; asthmatic; allergic; antibacterial; gene therapy; chromosome 19q13.2; autoimmune disease; multiple sclerosis; toxemia; juvenile diabetes; rheumatoid arthritis; pruritic urticarial papule; hypertension; Rh incompatibility; ss.			
XN	Homo sapiens.			
FH	Key	Location/Qualifiers		
FT	CDS	1..1932		
FT		/*tag= a		
FT		/partial		
FT		/product= "Preliminary CLASP-7 protein"		
FT		/note= "no start codon given"		
XN	WO200142295-A2.			
XN	14-JUN-2001.			
PF	13-DEC-2000; 2000WO-US34152.			
PR	13-DEC-1999; 99US-0170453.			
PR	14-JAN-2000; 2000US-0176195.			
PR	14-FEB-2000; 2000US-0182296.			
PR	11-APR-2000; 2000US-0196267.			
PR	11-APR-2000; 2000US-0196460.			
PR	11-APR-2000; 2000US-0196527.			
PR	11-APR-2000; 2000US-0196528.			
PR	11-APR-2000; 2000US-0547276.			
PR	13-OCT-2000; 2000US-0240503.			
PR	13-OCT-2000; 2000US-0240508.			
PR	13-OCT-2000; 2000US-0240539.			
PR	13-OCT-2000; 2000US-0240543.			
PA	(ARBO-) ARBOR VITA CORP.			
PI	Lu P, Garman JD, Candia AF;			
WI	UPI; 2001-381641/40.			
PSDB	P-PSDB; AAB99540.			
	Novel cadherin-like asymmetry protein-7 and polynucleotides encoding the polypeptide, useful for treating autoimmune disease, hypersensitivity, preventing transplant rejection by modulating immune			

QY	6161	atcataataaactgcgcagctctgctttaaagatttctaactaaaggtgtgaagatgccttaa	6220
Db	1685	atcacacaacaattgcgcctctgcttcaagacttctgcagaagaattgtgagatgcgtgc	1744
QY	6221	gaaaaaataagagcttaatttggccggttcaaaagagtatcaaaaggaattg--gggaa	6278
Db	1745	ggaaaaataagccctgatttggccgcgacacagaagagtaaccacgtgtgactgagcgca	1804
QY	6279	actatcttcgccttaaaagggccctacagccctagatcacagaagaagtcctcagttatcc	6338
Db	1805	actactgcgccttgcgggaagcctctgcagcccttctaccacgcgccttgcctcccaagctga	1864
QY	6339	aagccagtatctgtctgtcccttgcacacagagattccttcagtcgaatgagcttgcgaaa	6398
Db	1865	tggcaccacaccc-----accgcgcctcagaactccttgaacagagaagcaagtttccgaaag	1920
QY	6399	atgcatctctaacc 6412	
Db	1921	gcagacctctgagc 1934	

RESULT	10
AAS07381	
ID	AAS07381 standard; cDNA; 2014 BP.

AC AAS07381;

DT 26-SEP-2001 (first entry)

DE Human DNA associated with CLASP-5 #1.

KW Human; CLASP-5; cadherin-like asymmetry protein; immune gateway;

KW acquired immunodeficiency syndrome; AIDS; ss.

OS Homo sapiens.

	Key	Location/Qualifiers
FH	CDS	2..2014
FT		

```

12 /csg a
FT /product= "Protein as displayed in AAU04025"
FY /partial
FT /note= "No start codon"
FT /transl_except= (pos:1523..1525,aa:Arg)
FT /transl_except= (pos:1883..1885,aa:Ile)
FT

```

PN W0200142296-A2.

PD 14-JUN-2001.

PF 13-DEC-2000; 2000WO-US34163.

PR 13-DEC-1999; 99US-0170453.

PR 14-FEB-2000; 2000US-0182296.

PR 11-APR-2000; 2000US-0196460.

PR 11-APR-2000; 2000US-0196528.

PR 13-OCT-2000; 2000US-0240503.

PR 13-OCT-2000; 2000US-0240543.

XX

XX

PI Lu P, Garman JD, Candia AF;

DR WPI; 2001-367865/38.

DR P-PSDB; AAU04025.

XX CLASP-5 polynucleotides, proteins and antibodies are used to prevent or
PT treat a CLASP-5 mediated disease, such as an autoimmune disease e.g.
PT Rheumatoid arthritis -
XX
PS Disclosure; Fig 9A; 188pp; English.

Disclosure; Fig 9A; 188pp; English..

CC The sequence encodes a protein presented as AAU04025, included in the
CC specification which relates to a cadherin-like asymmetry protein,
CC CLASP-5, which is a transmembrane protein of the immune system involved
CC in the formation of the immune gateway. CLASP-5, polynucleotides encoding
CC it and an anti-CLASP-5 antibody are used to prevent or treat a CLASP-5
CC mediated disease, such as an autoimmune disease caused or exacerbated by
CC increased activity of TH1 cells. These diseases may include Addison's
CC disease, rheumatoid arthritis, dermatitis, Multiple Sclerosis, Systemic
CC Lupus Erythematosus and Autoimmune Thyroiditis, inflammatory conditions
CC (e.g. ischaemia-reperfusion) and responses, Leukaemia, acquired
CC immunodeficiency syndrome, AIDS, syphilis, typhoid, Burkitt's lymphoma,
CC Grave's disease and insulin dependent diabetes mellitis. CLASP-5 is used
CC for inhibiting an immune response in a cell (T cell or B cell) by
CC interfering with the expression of a CLASP-5 gene in the cell, the
CC ability of a CLASP-5 protein to bind to another cell or the ability of a
CC CLASP-5 protein to bind to another protein. The antibody binding CLASP-5
CC is used to inhibit an immune response is a subject. The polynucleotide is
CC used to detect CLASP-5 expression in cells and for diagnosis of
CC diseases and disorders associated with aberrant expression of CLASP-5.
CC Note: The present sequence is included in the specification but is
CC not mentioned anywhere in the specification.

Sequence 2014 BP; 565 A; 469 C; 514 G; 466 T; 0 other;

Query Match	12.3%;	Score 841.6;	DB 22;	Length 2014;
Best Local Similarity	66.2%;	Pred. No. 1.3e-201;		
Matches 1278;	Conservative	0;	Mismatches 639;	Indels 13;
				Gaps 4;

QY	4427	aagcaacctaatcatcttagatatacatlagagattgtgttcagaccgcttctgttaacgg	4486
Db	9	aacgagatttaatcatctctgtgatatgcaggaacaacattatccaggcgagctcggtcttgg	68

QY 4487 aatccaagagagcattcttggtygagtgtctaagaigtctactacacagcatgycctyta 4546
| | ||||| | | | | | | | | | | | |
Db 69 actgtaaagcacagcctgtcttggaagtgttcttgagggtgcctgttgaattctctgaactytg 128

QY 4547 accaaagtgcagtttatctacaacactgttctgtcacagagagccttggtttcaagt 4606
| | | | | | | | | | | | | | | | | |
Db 129 atcagagttaccacctactctactgaactgcttggacaacacttcgtgtctcatcgccaagt 188

QY 4607 ttctctgaactcttatttgaaagaagacagacagcagtgctgtcattatgcctcagcttc 4666

Db 189 ttggaagacttactcttcgaagaaggagtgtagaacagtggttcgcacctatgtcaccaagttcc 248

QY 4667 tccgacactgttagcaqtatgcatacagtcacaccccagtgtcccttacc 4726
| | ||||| ||||| | | ||| || | ||||| |||||
Db 249 tcgacacctgcagcagcagcatgatgtcaccccgagccaagctgtgccaccttac 308

QY 4727 tactaatgagcacaacttgygaattgygaataactttccagggttaaatgcaggtac 4786
| | | | | | | | | | | | | | | |
Db 309 tectcatgaggttcagtlttgygaccaccagtaatttggcaagagtaaagatgcaggtaa 368

QY 4787 caatgtcaactatacctcccttgtytggcacatatctcagaattttaatgaagaattcttaagac 4846
| | | | | | | | | | | | | | | | | | | | | | | |
Db 369 ccattgtcccctgycatcttttgttytgggaagagcacaccagaccttaatgaagagcaccttgagaa 428

QY 4847 gtcctcaaaagactatatattacatatgtctgaagaagatctygaattgagggaacaacat 4906
| | | | | | | | | | | | | | | |
Db 429 gatccttgaggaacaatttttgcctatttcagaagaagacacggccatgcagatgactcct 488

QY 4907 ttcttgatcaggtccaggatctggttttcaatctccatatgatctcttctgatactgta 4966
||| ||| | | ||| ||| | | ||| |||
Db 489 ttccccaccaggtggaggaacttctgtaatctgaaatagacattatagcacagcga 548

QY 4967 aaatgaagaaacaccagagatcctgaatatgttgattgatcctaattgtacagaattgcc 5026

Db	5430	A-----	5428
QY	2651	ccacaatgctagatctgcggtgagacctgcaagccttaattaaatcgttctcgaagcc	2710
Db	5429	-----	5428
QY	2711	ttagtaatagcaatccagatatatctgaggactcccagctccacagatgataaagttcgat	2770
Db	5427	cctttcaaaagcaaacagcagactaacatgcagctccagcacacagaga-----	5377
QY	2771	caatcatcgggagtaagggttagatcgctccaattccctgggttaaacactggtgtccaa	2830
Db	5378	-----CTTGCAATTAAAGCGAATCGTACACCATATATGATTAATGTTTGAAGAGTGCGCGA	5324
QY	2831	aagctgcccatggygagatcccaaccccaagtcaccaagtcagaatcaacacagctatgatac	2890
Db	5323	AGCTTGATCG-----CAAAAGTAAATATACATATTTGTAAAGGCTTTAAATA	5276
QY	2891	gaagttgtaatcgtatgtcttcgcacacagagagcgtcaagttcttacaaacattacg	2950
Db	5275	TAACTTTTCTTCCCTTTCAGAGCTGTCCATAGA-----TGTCCTTCACTCGATGGCTG	5225
QY	2951	gacgcttaccaactaaaagctttttcacgagagagctggtcttgcaagtggtgttttca	3010
Db	5224	CACGGGATGTCCAGAGTGGCGCTGCTTCACGAGAAGATTTAGCCTTGCAATTGGGTTGTGCAA	5165
QY	3011	gtggcagcgttcggaatacagctttgcacaagaagcctgtgtcttttgaatlaatgttaa	3070
Db	5164	GTGGAAGGCGAGCTGACTTAGCCATGTCAAAATTCGTGTTTATTGAGCTTATTGTCA	5105
QY	3071	agaagcattggtcaccaattatacttlaatgataaacttgaagctccaagaaagtcgtt	3130
Db	5104	AATCTATGATTGAACATTTGCAATTTGTTCAAAATACCTTTAAATGTTGCCCGCAACATCGAT	5045
QY	3131	ttccagaacgtttcatgatagacatltgcagctctctgtccagcacagatgtctagtataag	3190
Db	5044	TTCCACATCAGTTTAATGATGATCTATCCACCCTGGTGCACTTAGTTACACAACAAGTGG	4985
QY	3191	tttcacgatttcagaagcacagaaatggtltgagagactcaatacaagccttgcatctc	3250
Db	4984	TCGGCTACCAACAGCAACGAACTTAATTAGCGCAGTCATTAAATGCAAGTTTAGGCTTTT	4925
QY	3251	ttctcaatgactcgttctgtctgttatgacagagatgtgtttttagcctataaagtcct	3310
Db	4924	TTATATTGACATACTAAGCATCATGGATCGGGGATTTGTTTTCGGCCTCATCAAAACGT	4865
QY	3311	gctataaacagtgctctcaaaagccttactcaatccgaaatcccaagtgctgtctgtcct	3370
Db	4864	ACACAAAGGTTCTGATTTCTAAAAATGCTTCAATACG-----GATCTAATGAAT	4814
QY	3371	tgaagcgtgattttctcagaatcatctgcagtcatagcaactatgtttacataaacttac	3430
Db	4813	ATAAATATAGACTTTTGAAGAAATTGTGTAGTCAATGACATTTTGTAGCCTTAACCTTGC	4754
QY	3431	cctgcagcttacttactccacactgcatctccatcacacttctgtttcttctgcacaatctc	3490
Db	4753	CGTTTGGTACTTCATACACAATGGTAACAGCGCCATGACGCCACGCCAAGCACACAAGCT	4694
QY	3491	agagttctgattttctcagaatgtlacaagaaccaaaagattgcaatatgtttgtaattat	3550
Db	4693	CGAGCAACAGCCAAACTTCTTGCGTAAGTAAAAAACAATACACTTTAATTTTACTTAT	4634
QY	3551	ccgtgccttccgccaacagcaattattgycagga-----	3585
Db	4633	TTGTGGCTTCTTTTAAAGGGATCAGTAGAGAGAGCCCTTCACGCCGACCTTAGTCAAGA	4574
QY	3586	-----ctgtgttaacagagctggctgtcatatttagaccctgatgc	3626
Db	4573	ATTCCCTCAGCAACATTTCTAGTCGAGACTTGTTCGAGCGACTTAGCAGCAGTAATGGA	4514
QY	3627	tgaagcactgttttgatgtcataagaagtcataataatgtttacacaatttactctccag	3686

Db	4513	GGTGCCAAACCCCAATTACATGGAAAAAGCAATAAAGGTATTTCCGAACCTGATGCATC	4454
QY	3687	tcacgaactcagaccgcgcgtactctgaacctcagataaagctcgagtgccatgtgta	3746
Db	4453	TCACGATTTGGACGGCGGTTATAGCGAAACTGATGCACGTCCAAAGGGTGGCTTTATA	4394
QY	3747	tctacctgtatgtgatatcatcaltgynaacctgtacctcagctgtatgatlttacgaaac	3806
Db	4393	TATACCGCTACTTTCTATAGTGATGGATTGCATTCCTCAGTTGCATCAGCAGCGCTTGA	4334
QY	3807	tcacaatcaacgaggaagaccaaattgtatagccactgatgtattatgaagtgaagcgg	3866
Db	4333	GCAGGATCGTCTGCAGCAAAATCGGTCAACT-----TGAGACTATCAAGGGCCACATCA	4280
QY	3867	aagtatgataagccagagaccggttgcacatgycaatgcgagggacatcggtccctcaactaac	3926
Db	4279	AACTATTACAGCATCAACCAATAATCCGGAAGTGGCA-----TTTGGCATATTCGG	4230
QY	3927	aagccctgycagtttccctcctcaacgltcaacgagtgycagggcaacacactacatttcagc	3986
Db	4229	GAAGTGTCCTTACTCCTTACTTGAATGAACAAGTTAAACAAATTGCCCTCAGCTCGG	4170
QY	3987	agaatcaagtcgaagccctttgatctgtctactttgggttctcaaaatgfcagatgaaac	4046
Db	4169	AGAAAT--ACCCGCCATTTATTTGGTGTGTTTCTGTGGGTCTAAAAAATCTGGAGCGCAA	4112
QY	4047	agttctacagaagtggtttacagatctctcagtccttgcagctaaaccgctattagatct	4106
Db	4111	CGTTTATATCGCTGGCTTATGGATTTAAGTCCGCATAGAGTGCACCAATGCTTCAAGG	4052
QY	4107	gcttatctctgtgtctcttgcctttagtataaagggaaaaaagtgtttgaaagaaatgaa	4166
Db	4051	TGTTAAAGTTTGCTTGAANAATTTGAGTACACAGCACAATA-----AA	4007
QY	4167	tagcttgaccttlaagaaaatcaaaagacatgagagcaaaagcttgaagaagctattcttg	4226
Db	4006	CGTGCCACCCCTTAACGAAACGAATACACAAAGTTTAAAGAAACTGGATCTACTGATGT	3947
QY	4227	gagcatagtgccagcgcaagaataatggtacgycgaagccgagagacagctcgagagaagccc	4286
Db	3946	TAAAGAAAAGCTGGAAGATGTATTAGAGGCACAAATTCAGCACGTTATGATTTAATAAA	3887
QY	4287	atctggaagtgcccttgggaagtcaagaaaatttggagtygaggaagaatatgactcactg	4346
Db	3886	TCGTGCAAAAGATCGTAATTCAACAGAAAATTCGGAATGGAGGAAGACCCAGATGCCATA	3827
QY	4347	gcg---tcaaaacacagagaagaagcttgcacaatcaagaagcagagattgaacacgaagcact	4403
Db	3826	TCGCTCTCAATATGCGGATGCTGTGGAAAAAGTGAGCATGAACCTGAATTAAGTCATTT	3767
QY	4404	gattgatgaaaacctgctacagaaagcaaacctaatcatatttagatacatlagaatgt	4463
Db	3766	CATCGAAGGCTTTTGGCAACTGAAGTTGCACCTTGTACTTCTCGATACCCCTTGAATAAT	3707
QY	4464	tgttcagaccglttctgtaacggaatccaaagagagcattcttggtygagtygctaaaagt	4523
Db	3706	TGTTCA-----TGCTGCTGCCAACCTTTATCACAACTTCTTGGAACTGTGCTAAAGCT	3653
QY	4524	gctactacacagcatgycctgtaaccaaatgtcagtttatctacacactgtttgtctac	4583
Db	3652	GCTTCTTCATTCCTTGTGCCCGGAATCAATCTGTTTGGCCCTGCAAAACCTGTTTGCCCTC	3593
QY	4584	acagagagccctggtttccaagtttccctgaacctctatttgaagaagagacagagcagtg	4643
Db	3592	CCAACGCGGCTTAATCTTTAAATTTCCAAACTTAATTTGATGACGAGACCGCATATATG	3533
QY	4644	tgctgatttatgcttcaggcttctccgacactgtagcagtlagcatcggtacataacggtc	4703
Db	3532	TGCGGATTTGTGTAAATTTTACTGAACAATGTGGGTCCCTATTTGCCAGGAATACGATC	3473
QY	4704	acaccccaagtgcctcccttacctactaatgagycaaaactttgagattgggaat-----	4758
Db	3472	ACAGGCAGCTGCATCCTATATTTACTAATATGAGCAAAATTTGAAATTTGGAATGTAAAG	3413

QY 4759 -----aact 4762
Db 3412 TTTTATTATTGTTAGGTTTATTACATAAATTATGGTCACTTATTCGTATAGAACT 1111
QY 4763 ttgccaggttaaaatgcaaggtaccaatgtcaatacctccttggtggtgacacatctcaga 4822
Db 3352 TTGCCCGAGTTAAGATGCAAGTGACGATGTCTTTAAGCTCCCTAGTTGGAACAAGTTCGG 3293
QY 4823 atttaataagaataatcctaagaagctctcctaagaactataatgacatatgtcgaagaag 4882
Db 3292 TCTTTAGTGAGCAATCCTTACGCCGCTGCGCTTAAACCGTCCGTGTTATGCTGAATCCG 3233
QY 4883 atctggaattgagsgaacaacaacattcctgatacaggtccagatctggtttcaatctcc 4942
Db 3232 ACTCCGACCTGCAGAGTAACCTCTTTCCGTGAACAAGTACAAAGATTGCTTTTAAATCTGC 3173
QY 4943 atatgattccttctgatactgtgaanaatgaaggaaacccaaggagatcctgnaatgttga 5002
Db 3172 ATATGATCCTGTCAAGTAACCTGTTAAATGAAGAGTATCAGGAAGACCCAGAATGTTGC 3113
QY 5003 ttgatctaattgtacagaattgccaaggttaccagacctctcaga---gagattgacct 5059
Db 3112 TTGACCTCATGAATCGTATTTGCCAAGGATACCAAAATAACCTGATCTAGACTGACTT 3053
QY 5060 ggttgacagaacatgycagggcaagcactcagaaagaaagaatcatgtcgaagctgcacagt 5119
Db 3052 GGTGGAAAAATATGGCTTAAATAAACACCGCGAGCGCAAAATCACACGGGAAGCCATGT 2993
QY 5120 gtctagtcacactcagcagcactgtgtcgaataattgagcatgtcgaagccggaanat 5179
Db 2992 GTTATGTACATGCTGCTTCTTTAGTTCTGAATATCTTAGCATGTTGAGTCACAAAAAC 2933
QY 5180 atcttcctgtggatgtgtaacaattcagaataatccatctaattgattagaagaatctg 5239
Db 2932 AATTGCCCTGTGAGCTGTAAAGTTTCAACGAATTTCTCCAACACACATAATGAGTCGG 2873
QY 5240 cggctcagatgatgtgtatctccagatgaagaaggtatctgctctgnaaataacttta 5299
Db 2872 CCGTATCGGATGATGTGTAAGTCCCGGCGAAGATGTATCTGCTAGGAATCATTTTCA 2813
QY 5300 ctgagtcagacttgtgtgatacttgaacaagcagctgtcctcctctatgtgtgca 5359
Db 2812 CTGAACCTGGGTGAAGGCTTGCTGGAAGAAGCCCTCCAATTCTTTCAAGTTGCTGGCA 2753
QY 5360 tgtatgaagcaggttaattgaagtttacaagaagtaactatctcattcatgaagtaatcggg 5419
Db 2752 TGTATGAAGCAATGAACGAAGTGTACAAAATTCTAATACCATATATGCGAGGCTTAACAGAG 2693
QY 5420 atgcaaaagaactatccacaattcatgtgaactcaagaagcatt----- 5465
Db 2692 AATTTCAAAAAGCTAAGCAAGTTCATGGCAAAATGCGAAGGCATTTAATCGAATATCCC 2633
QY 5466 -----cagcaaaatgttcatcagagtact 5490
Db 2632 AACTACAGGTAACAATATTTGTAATTTTACCACGGAATAATATATACATATTTATAAA 2573
QY 5491 ggcctggggagcgatgttttgcaacctatttcgtgtgtttttatgnaaccaagttcggg 5550
Db 2572 CAGGTAAGAGAGTTTTGGACATACACTTTCGTGTGGCTTCTATGCGGGAATAATTGGG 2513
QY 5551 gatttgatgaacaagaatttgtttacaaggagcctgcaataaaccaaacttgcaagagata 5610
Db 2512 GACTTGGATCAGCAGGAATTCATTTATAAGAGCCACACATTGACGAAGTTGCCCGAATA 2453
QY 5611 tct----- 5613
Db 2452 TTTAGTCGGCTTCAGGTATATATTGCAAAATTGGAATAATAGAACTAATCAATTTTGTTP 2393
QY 5614 -cacagattgaggtatttaaggaagaagatttgagagagatgtgtgaagtaatcaa 5672
Db 2392 CAACATACGTTAGAACTTTTACACTGAAACGATTCGACCGGACTCTGTGCATATCATTTAA 2333

QY 5673 agactctaactcctgtagacaagtgtaaatagatcctaacaaggcatatattcaagattac 5732
Db 2332 AGATTCCAATACCGTTGATATTAATAGCTTGGATCCGATPAAGGCTTACATTCAAATTTAC 2273
QY 5733 ctatgtgagccatacttgcacacatatgagatgaagagacgaatcacatttcgacaa 5792
Db 2272 TTATGTTGAACCCCTACTTGTGAACATATGAATGCGTCATCGTGACACATACCTTGAGCG 2213
QY 5793 aaattacaactctcgt----- 5808
Db 2212 GAATTTCAAATATAAGTATGATATGAATTAACAGATTAATTAAATCGAAATTTAAATA 2153
QY 5809 -----cgattcatgtactgttacaccccttacttag 5839
Db 2152 TAATTTAACATTTTGGTTATATGACGAACGTTTATATATATGCCACGCCCTTTACTAATA 2093
QY 5840 atggccgtgcccatggygaacttcatgaacaattcaaaaggaagaccattctgactagt 5899
Db 2092 ATGGAAGGCACATGCGCAACTAAATGAGCAGTGTAAACGGAACGATTTTGACCAACGG 2033
QY 5900 ctcatgcttccctlatataaacaaggytcatgtcactcataaagaagatcatct 5959
Db 2032 CGAATCACTTTCATACGTAAAACTCGCATTATGTTAATTAGCCGACAGCAATTAAGTCC 1973
QY 5960 taacaccaattgaagltgtatg----- 5983
Db 1972 TGAACCCGATTTGAAGTGCAATTCGAAGTAAATTATGTTGTAATAATCCCAAAAGTTAATA 1913
QY 5984 -----aggacatgcagaaaagacacaggaagltgacatttgcaaacac 6025
Db 1912 AATTAAACGTTATTGGCTAGGATTAATTCAAAAAAACATTTGGAATTTGGCAGCTGCCACGA 1853
QY 6026 atcagatcccgacagaccccaaatgcttcagatgtgtactccaggaatctgtlagcacca 6085
Db 1852 ATCAAGAGCCACGACAGACCCCAAAATATTGCAAAATGGTACTGCAGGGATGTATTGGAACCA 1793
QY 6086 cagtgaaatcaggggccttgaagattgccaggttcttctgtcgaataactagtagacc 6145
Db 1792 CTGTTAACCAAGACCGATGGAATGGCGAGTGTGTCTTCCAAATTTATCCGACGGAA 1733
QY 6146 caa--agctcttcaagacatcataataaactgcgactcgtcttaagaatttactaaaa 6202
Db 1732 CAACGTACCGACAAACACCAACAAACTTCGTTGTGCTTTCGCGAGTTTCAAAAGC 1673
QY 6203 ggtgtgaagatgccttaagaaaaataagagcttaattggccggttcaaaagagatlc 6262
Db 1672 GTTGTGCTGATGCTTTGAAAAAGAAATCGCAATCTAATACTTTCAGATCAAAAAAGTATACC 1613
QY 6263 aaaggaattgg 6274
Db 1612 AACGAGAACTGG 1601
RESULT 12
AAH99572
ID AAH99572 standard; cDNA; 1652 BP.
XX
AC AAH99572;
XX
DT 16-OCT-2001 (first entry)
XX
DE Human protein encoding cDNA sequence SEQ ID NO:407.
KW Human; cancer; ulcer; HIV infection; human immunodeficiency virus;
KW antinflammatory; antirheumatic; antiarthritic; immunosuppressive;
KW antibacterial; endocrine; cardiant; central nervous system; virucide;
KW anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; anaemia;
KW antiagregant; haemostatic; vulnerary; antituber; osteopathic; eczema;
KW dermatological; antiallergic; antiaslhmatic; antidiabetic; cyostatic;
KW neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;
KW immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;
KW antianaphylactic; rheumatoid arthritis; septic shock; pancreatitis;
KW cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;

KW genetic disease; haematopoietic disorder; platelet disorder; asthma;
 KW thrombocytopenia; osteoporosis; severe combined immunodeficiency;
 KW allergic rhinitis; diabetes; multiple sclerosis; depression;
 KW Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;
 KW neurological disorder; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200153455-A2.
 XX
 PD 26-JUL-2001.
 XX
 PF 22-DEC-2000; 2000WO-US35017.
 XX
 PR 23-DEC-1999; 99US-0471275.
 PR 21-JAN-2000; 2000US-0488725.
 PR 25-APR-2000; 2000US-0552317.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Drmanac RT;
 XX
 DR WPI; 2001-457603/49.
 DR P-PSDB; AAM25631.
 XX
 PT Isolated human polynucleotides encoding polypeptides, useful for the
 PT treatment and diagnosis of e.g. cancer, ulcers and HIV infection -
 XX
 PS Claim 1; Page 505; 1217pp; English.

CC AAH99166 to AAH99904 encode the human proteins given in AAM25225 to
CC AAM25963. The proteins can have activities based on the tissues and
CC cells they are expressed in, such as: antiinflammatory; antirheumatic;
CC antiarthritic; immunosuppressive; antibacterial; endocrine; cardiant;
CC central nervous system; virucide; anti-HIV; fungicide; antimutagen;
CC cardiovascular; antianaemic; antiagregant; haemostatic; vulnery;
CC antiulcer; osteopathic; dermatological; antiallergic; antiasthmatic;
CC antidiabetic; cytotstatic; neuroprotective; antidepressant; nootropic;
CC antiparkinsonian; and immunostimulant. The proteins and polynucleotides
CC encoding them can be used in gene therapy, antisense therapy and vaccine
CC production. The proteins and polynucleotides are useful for screening for
CC agonists or antagonists of a protein and for the treatment and diagnosis
CC of disorders associated with the activity of a protein e.g. inflammation,
CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,
CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal
CC infections, autoimmunity, genetic diseases, haematopoietic disorders,
CC anaemia, platelet disorders, thrombocytopaenia, wounds, burns, ulcers,
CC osteoporosis, severe combined immunodeficiency, eczema, allergic
CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,
CC Alzheimer's disease, Parkinson's disease, neurodegenerative and
CC neurological disorders.

Query Match	8.3%;	Score 565;	DB 22;	Length 1652;
Best Local Similarity	67.1%;	Pred. No. 6.7e-132;		
Matches 862; Conservative	0;	Mismatches 415;	Indels 8;	Gaps 4;

[illegible]

Db	183	tggtgaggttgctgtggaacagcgccggtactcttcaaccatggygcgggtctacgagcg	242
QY	5372	ttaatgaagtttcaaaagtlacttatctctatctcatgaagctaacgcggaatgcaagaac	5431
Db	243	tgaatgaggtctacaagaacctcatcccatcctctggaagccacccgtgactacaagaagc	302
QY	5432	tatccaacatctcatgtaaacctccaagaagcatctcagccaacattgttcatcagagtactg	5491
Db	303	tgccgcggtgcacgcgcaacctgaaggagcgcttcaccaagatcatgcaccagattccg	362
QY	5492	gctgggagcggaatgtttgccaacctatttcgtgtgtgttttatgtgaaccaagttcggg	5551
Db	363	gctgggagcgcggtgttctcggaacgtatttcgcggtggtgcttcaagcgcccaattcgtg	422
QY	5552	atttggatgaacaagaatttgttttaacaaggagcgctgcataaaccaacctgcaagatat	5611
Db	423	acctggaatgagcaggaagtgtgtatacaaggagccatcgatcaagagctgycagaagatct	482
QY	5612	ctcaacgaattggaaggatttaccggaagaagatttggagagatgtgttgaagtaatca	5671
Db	483	cacaccggtctggaagagttctacaacgagagatttggcgaacgctgcgttgagattatca	542
QY	5672	aagactctaactcctgttagacaagtgtaaatlagatcctaacaaggcatatatcagatta	5731
Db	543	aagactctaaccctgttgacaagttccaagcttgactcaacaaaagcctacatccagatca	602
QY	5732	cctatgtgagccatactttgacacataatgagatgaaggacgaatcacctatttcgaca	5791
Db	603	cgtatgtggaaccgtactcttgatacctaagcgtcaagacgagccggtgacctactttgacc	662
QY	5792	aaaattacaactcttcgtcgatctcatgtactgttacaccccttactttagatgycggtgcc	5851
Db	663	gcaactatgggtcttcgcacatctcctgttctgcacgcccgttcaacgccgatatgycgcac	722
QY	5852	atggggaaacttcatagaacaatccaagaagagaccatcttcgaactacgtctcatgcttcc	5911
Db	723	acggggagctgcccgaacacacaagcgtlaagacgttgtcagacccaagccacttcc	782
QY	5912	cttatattaaacaaggtccaatgtccaactataaagaagatcat-cttaacaccaatt	5970
Db	783	ctacatcaagactcgcatacgtgtgtgtgccaccggaagagacggtgtcagycgccagt	842
QY	5971	gaagttgtctattgagagacatgcagaaaaagacacaggaattggcatttgcacacatcag	6030
Db	843	gaggtgccaatcgagagacatgcagaaagaagacacggaagctggccttggccaccgagcag	902
QY	6031	gatcccgcaagcccccaaatgtcttcagatggttactccaaggatctgttaggccaccaagt	6090
Db	903	gaccacacagatgtctaagatgtacagatgtgtccttcaaggctctgtagggcccaacctg	962
QY	6091	aatcagggccttgggaagtgtgccaggttttctgtctgtaaatatacctagtgaaccaag	6150
Db	963	aaccaggtccctgtgaggtggcccaagtgtttttagcagagatcccggaagaccccaag	1022
QY	6151	ctcttcagacatcataataaactgcgactcgtgctttaaga-ttctaataaagtgta	6209
Db	1023	ctcttcggaatcacacaacaaattgcggtcgtgtcttcaagactctgcacaagaatgtga	1082
QY	6210	agatgccttaagaaaaataagagcttaatttgggcggttcaaaagagtatcaaaagga	6269
Db	1083	ggatgcgctgcggaaaaaataagccctgatttgggcggaaccagaagagattaccaccgtga	1142
QY	6270	attg--ggaaactatcttcgccttaagagcgccctacagccctagatcacagaagaatcc	6327
Db	1143	gctggagcgcaactacttcgcgcctgcgggaggtcttcagccctgttaccceagcgct	1202
QY	6328	ctcagttatccaagccaagtattgtcttccctgcaccaagagattccttcaagtcgaatga	6387
Db	1203	gccccagctgattgcaaccacccc---accggcctcaggaactccttgaacagaagcaa	1258
QY	6388	gcttcgcgaanaatgatactctaaac	6412
Db	1259	gtttccgaaaagcagaccttgaac	1283

RESULT 13
AAI61331
ID AAI61331 standard; cDNA; 2427 BP.
XX AAI61331;
AC
XX 22-OCT-2001 (first entry)
DT
XX Human polynucleotide SEQ ID NO 5320.
DE
XX
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokine; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia; ss.
XX
XX Homo sapiens.
OS
XX
XX WO200153312-A1.
PN
XX
XX 26-JUL-2001.
PD
XX
XX 26-DEC-2000; 2000WO-US34263.
PF
XX
XX 21-JAN-2000; 2000US-0488725.
PR
XX 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX
XX (HYSE-) HYSEQ INC.
PA
XX
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
DR WPI: 2001-442253/47.
DR P-PSDB; AAMA2175.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
PS
PS Claim 1; SEQ ID NO 5320; 10078bp; English.
XX
CC The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cytosstatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
SQ Sequence 2427 BP; 716 A; 508 C; 524 G; 679 T; 0 other;

Query Match 6.7%; Score 459.8; DB 22; Length 2427;
Best Local Similarity 68.8%; Pred. No. 2.8e-105;
Matches 678; Conservative 0; Mismatches 297; Indels 10; Gaps 3;

QY 5369 cagttaatgaagtttacaagaagtaatttccatttcatgaagctaatacggatgcaaga 5428
DB 4 cagttaatgaagtttacaagaagtaatttccatttcatgaagctaatacggatgcaaga 63
QY 5429 aactatccacaattcatgtaaaccttcaagaagcattcagcaaaattgttcatcagagta 5488
DB 64 agctgacactcaactcacaagcaagctgcagagagccttcgacaagcatcgtttaacaagatc 123
QY 5489 ctggtggaagcggatgtttggaacctatttcgtgtgtgttttatggaaccaagttcg 5548
DB 124 ata-----agagaatgtttgaaccttccagttcgagttgttcttltgataccaatttg 177
QY 5549 gggatttggaagacaagaatttgttacaagagagcctgcacaataccaacttcgacaga 5608
DB 178 gggatttggaagacaagaatttgttacaagagagcctgcacaataccaacttcgacaga 237
QY 5609 tatctcacagatttgagggattttacggagaagaatttggagaggtgtgtgaagtaa 5668
DB 238 tctcacatagactagagggcatttltatgtlcaatgttltgtgcagaatttgtgaagtga 297
QY 5669 tcaagactctaactcctgttagacaagtgtaaattagatccctaacaaggcataattcaga 5728
DB 298 ttaagactccaactcctgttgacaaaaccaagtttgatccctaacaaggcctacatacaga 357
QY 5729 ttacctatgttgagcccatcatttgacacatatgagtgaagacagaatcacctattcg 5788
DB 358 tcaactltgttgagcccatcatttgatgagtatgagtgaagacaggtgtcacatacttg 417
QY 5789 acaaaaattacaatcttcgtcgattcagtactgttacacccttactttagatggcgtg 5848
DB 418 agaagaatttcaacctccggaggttcattgttacaccaccctgtcaacctgtgagggcgcg 477
QY 5849 cccatggggaacttcatgaacaattcaaaaggaagaccattctgactactgctcatgct 5908
DB 478 ctcggggagagcgtcatgagcagtagcacagaagaaacacagctccgtgacactatgcagcct 537
QY 5909 ttccctatatataaacaaggggtcaatgttcaactcataaagaagagatcatcttaaccaca 5968
DB 538 tcccctacatcaagaaccagagatcagcgttccatccagaaggaagagttgttltgacaccga 597
QY 5969 ttgaagttgctattgaggaacatgcagaaaaagacacaggaagttggcatttgcaacacatc 6028
DB 598 ttgaagttgcatgtgaagacatgaagaagaagaacctgcagtttagcagttgccaataacc 657
QY 6029 aggatccgcagacccccaaaatgtctcagatgtaactccagggatctgtagcaccacag 6088
DB 658 aggaagccgctgtgacaaagatgtctcagatgtgtgtgcaaggtctgtggagactatg 717
QY 6089 tgaatcagggccttttgaagttgcccaggttttctgtctgaaataacctagtgaccaca 6148
DB 718 taatcagggacccaactggaagtagcccaagtgtttltgtgtaattcctgtgatccaa 777
QY 6149 agctcttcagacatcataataaactgcgactctgctttaaagatttactaaagtggtg 6208
DB 778 aactctatcgacatcacacaagaagttgagttatgtctttaaagattcatcatgagatgtg 837
QY 6209 aagatgccttaagaaaaataaagagcttaattggccggttcaaaaaggagtatacaaggg 6268
DB 838 gtgaagctgtagaaaaaacaagcgttcatcacgcgcagcagcagaggggaatatcagcagg 897
QY 6269 aattgggga--aactatcttcgcctttaaagagccctacagccctagatc--acagaag 6324
DB 898 aactcaaaaagaactataacaagctaaagaagaaacctcagcgccaatgtatcgagcggaaaa 957
QY 6325 tccctcagttatccaagccagtatt 6349
DB 958 ttccagaaactgtacaagccaataatt 982

RESULT 14
AAI59545
ID AAI59545 standard; cDNA; 2433 BP.

XX	AAI59545;	
AC		
XX	22-OCT-2001	(first entry)
DT		
XX		
DE	Human polynucleotide SEQ ID NO 1748.	
XX		
KW	Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;	
KW	peripheral nervous system; neuropathy; central nervous system; CNS;	
KW	Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;	
KW	amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;	
KW	chemokinetic; thrombolytic; drug screening; arthritis; inflammation;	
KW	leukaemia; ss.	
XX		
OS	Homo sapiens.	
XX		
PN	WO200153312-A1.	
XX		
PD	26-JUL-2001.	
XX		
PF	26-DEC-2000; 2000WO-US34263.	
XX		
PR	21-JAN-2000; 2000US-0488725.	
PR	25-APR-2000; 2000US-0552317.	
PR	09-JUL-2000; 2000US-0598042.	
PR	19-JUL-2000; 2000US-0620312.	
PR	03-AUG-2000; 2000US-0653450.	
PR	14-SEP-2000; 2000US-0662191.	
PR	19-OCT-2000; 2000US-0693036.	
PR	29-NOV-2000; 2000US-0727344.	
XX		
PA	(HYSE-) HYSEQ INC.	
XX		
PI	Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;	
PI	Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;	
PI	Zhao QA, Zhou P, Goodrich R, Drmanac RT;	
XX		
XX	WPI; 2001-442253/47.	
DR	P-PSDB; AAM40389.	
PT		
PT	Novel nucleic acids and polypeptides, useful for treating disorders	
PT	such as central nervous system injuries -	
XX		
PS	Claim 1; SEQ ID NO 1748; 10078bp; English.	
XX		
CC	The invention relates to human nucleic acids (AAI57798-AAI61369) and	
CC	the encoded polypeptides (AAM38642-AAM42213) with nootropic,	
CC	immunosuppressant and cytostatic activity. The polynucleotides are useful	
CC	in gene therapy. A composition containing a polypeptide or polynucleotide	
CC	of the invention may be used to treat diseases of the peripheral nervous	
CC	system, such as peripheral nervous injuries, peripheral neuropathy and	
CC	localised neuropathies and central nervous system diseases, such as	
CC	Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic	
CC	lateral sclerosis, and Shy-Drager Syndrome. Other uses include the	
CC	utilisation of the activities such as: Immune system suppression,	
CC	Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic	
CC	and thrombolytic activity, cancer diagnosis and therapy, drug screening,	
CC	assays for receptor activity, arthritis and inflammation, leukaemias and	
CC	C.N.S disorders.	
CC	Note: The sequence data for this patent did not form part of the printed	
CC	specification.	
XX		
XX	Sequence 2433 BP; 714 A; 505 C; 528 G; 686 T; 0 other;	

Query Match	6.48;	Score 434.4;	DB 22;	Length 2433;
Best Local Similarity	70.8%;	Pred. No. 7e-99;		
Matches 606; Conservative	0;	Mismatches 246;	Indels 4;	Gaps 2;

OY 5498 agcggatgtttgacactatttcgttgtgttlttatygaaaccaagtgcgggatcttg 5557
 || | ||||| | | | | | | | | | | | | | | | |
Db 139 agagaatgttttgaacctacttcocgagttgttctcttgygatccaaatttggggatcttg 198

[illegible]

RESULT	15
AAS62625	
ID	AAS62625 standard; cDNA; 1834 BP.
XX	
AC	AAS62625;
XX	
DT	14-FEB-2002 (first entry)
XX	
DE	cdna sequence #412 encoding novel human secreted protein.
XX	
KW	Human secreted protein; hyperproliferative disorder; autoimmune disorder;
KW	immune deficiency disorder; blood disorder; inflammatory disorder;
KW	infectious disorder; gene therapy; antimicrobial; hepatotropic;
KW	immunosuppressive; antirheumatic; ss.
XX	
OS	Homo sapiens.

XX WO200177291-A2.
PN
XX 18-OCT-2001.
PD
XX 29-MAR-2001; 2001WO-US10485.
PF
XX 06-APR-2000; 2000US-195604P.
PR
XX (GEMY) GENETICS INST INC.
PA
XX Wong GG, Clark HF, Fechtel K, Agostino MJ, Howes SH, Resnick RJ;
PI Gulukota K, Graham JR;
PI
XX WPI; 2002-010900/01.
DR
XX
PT New polynucleotides encoding secreted proteins useful for treating e.g.
PT asthma, HIV and Crohn's disease -
PS
XX Claim 1; Page 292-293; 391pp; English.
XX
CC The present invention relates to the isolation of novel cDNA sequences
CC which encode human secreted proteins. The cDNA sequences have been
CC derived from a variety of human tissues. The invention also provides
CC a method for producing proteins from these polynucleotide sequences.
CC The proteins are useful for identifying compounds that modulate their
CC activity and production, and the cell is also useful for identifying
CC compounds that modulate expression of the polynucleotide sequences
CC encoding the secreted proteins. The sequences of the invention are
CC useful for treating diseases such as hyperproliferative disorders
CC (e.g. cancer), immune deficiency disorders (e.g. severe combined
CC immunodeficiency (SCID)), autoimmune disorders (e.g. multiple
CC sclerosis), blood disorders (e.g. thrombocytopenia), inflammatory
CC disorders (e.g. arthritis) and infectious disorders (e.g. hepatitis).
CC The polynucleotide sequences of the invention are also useful in gene
CC therapy. AAS62214-AAS62838 represent the cDNA sequences of the
CC invention that encode for novel human secreted proteins.
XX
SQ Sequence 1834 BP; 502 A; 442 C; 480 G; 410 T; 0 other;

Query Match 5.4%; Score 370; DB 24; Length 1834;
Best Local Similarity 55.0%; Pred. No. 1.1e-82;
Matches 843; Conservative 0; Mismatches 670; Indels 21; Gaps 5;

QY 173 ttaccgaagcagtagatccagtgattggaagattaccctcatctcatccttgcgtg 232
Db 20 ttatgacctgtgagccagtgacttgaaagactctgtatgcacaacctgaacagcc 79
QY 233 tggattctggccttaccggaattgattgaatttcctccagatgatatgaagtgttt 292
Db 80 tggatgtgcagcttgcccaaggagctcggggacttccactgatgacgacttggacgtgtgt 139
QY 293 atagtcctcgagactgcagaactctgtttcagctgttacctgtaagaaag---tgaatg 349
Db 140 tcacgacaaagaaatgtagacttgcagccctcttgcggaggaagggttgaactg 199
QY 350 atccacatgttagagactgtataagaagttatacagaagactgggcaattgtcatcagaa 409
Db 200 accctcatgtccgggactgtgttcagactcatccgtgtgagtgttaatcgtgaaccgga 259
QY 410 aatatcataaatggygaacaggaatttaacccaatatcatattagataaacaagaagagc 469
Db 260 aaaccaaagaaagtcagaagaatctgtgtctttaaaaaagactggaatcgcgaaaagatttc 319
QY 470 aaaaaggttgcgcaaaacaagtttttgaatctgtagaagctccagatgccaacagctacc 529
Db 320 acaagaagcttcgcaaaacagacgttttgatcgga-----aaccttggagltgcagtgaac 373
QY 530 aggatgatcaagatgaccttaaaagacgttcaatgtcaatagatgataccccaagggtta 589
Db 374 ccgctgtcagcgagggcccccgcacttaaacgtgtgtgcgacgtgtctgggaaaggcc 433

QY 590 gctgggcctgtagtatcttgcacttgaaaaaattcacttccttgatgcttgcctccaatt 649
Db 434 ccgtcactgctgtgactttagacctccgcagcctgcagccttgacaagcggtagaanaacc 493
QY 650 tacttgatcgaaactccaatatgaagaatagaccggtcagaatgatgaccaagaagaatcaa 709
Db 494 tcctgcagcaagtgtgctgcgaggaacttggaaagcagaacgagggagcccgagagacca 553
QY 710 accgtcacaaagaacttttgccttgcatacccatcaccagatgggaaaccaatagaac 769
Db 554 atagcgagcgcgagctcttgccttaccatcagtcagcgaggaagatgctgtggaaa 613
QY 770 ggcttagtgttccctgatatacccaagaacatttgcgtcaaaagacttctgtaaaatgct 829
Db 614 taagtcagtagtaagaatgtcccaagaacacctgggcaacagaataattgltcaagltgc 673
QY 830 tatcactcaagtttgaaatltgaaatlgaacccatttgcgaagtttgcttatatgatg 889
Db 674 tgaccttgaagltcagalttgaaatltgagccctgttgcagacattgcctctaagatg 733
QY 890 tcaagaaagaagaaagatttcagaanaactttatttgaccttaattctgagcagatga 949
Db 734 ttaagaagaagaaagatctcagaanaatttccactgtgaccttgaaactctgaccagltta 793
QY 950 aaggtgtgtacgtccacatgtlaccacctgtgcattactlaccctggcgaagatcagcaa 1009
Db 794 aagatttctgcgagctcacacgccttcagtgggccgcatcaagtcaaggcgagatctgcag 853
QY 1010 ttttctatcactatccctcccaagatgttttctgtataaagctagaanaagttcc 1069
Db 854 tcttctcagtcacctaccctcctcctcagacatctactctgtagtcagaagtgaanaagttcc 913
QY 1070 tacagcaaggagacattggagagtggtgcagaaacctatatatgatattccaagaagcagatg 1129
Db 914 tgcagcagggagagatttggagactgtgcagagccctaacacggttatccaagaagtgatg 973
QY 1130 ccaaccaagaataaagaaanaacttggagaacttgaagaagtcgaagcagatcagtttgcacaa 1189
Db 974 gtggaagaagtaagaanaaagattgaaaaaactccaagctgaatccttctgcagc 1033
QY 1190 gacttgggaatatatgcacatgccttltgtgcgtgactgcaatcccatthaatgaatatgtta 1249
Db 1034 gtttgggaatataccgagatgccttltgtgcgttggcaccataagaacttaacagcttct-- 1091
QY 1250 gcaagtgtggagtttggaaagagatctacagaagtagaatacagtaactggaagaa 1309
Db 1092 -caatgtctccacccttgagagggaagttaacctgattgga-----ctctgtgttggga 1144
QY 1310 aaggtctctgttcagaggaagggaattctatgtattgttgcagacagatcacttgaagga 1369
Db 1145 gaagctcagtggttgaacggaaggacatttggccaatctagaaggcttcc--tgaagag 1201
QY 1370 caacaagltgagatgatgtcttgaacttgacagacttgcaccagctactctcacagtga 1429
Db 1202 ccctctccttggaggaanaatggggttggatccaacttcaaacctccaactctgagcgta 1261
QY 1430 caaattttttaaagcaggaaggagacccgcttaagtgtatgaagatcctacaatctcctg 1489
Db 1262 gcagcttlttcaagcaggaaggagatcgccttaagcagatgaagacttatccaagttttag 1321
QY 1490 ctgatatgagaagccactctctgtcttaccgagactaagacctaatacagctcagctca 1549
Db 1322 ctgactacaagaatcatcatccttaccagagcaggtcaagtcaattccagcgcttgtaa 1381
QY 1550 agatagacatttctcccgacactgaaatccccaattatlgcctaactccgagctgtctc 1609
Db 1382 gactggagatttctacagctccagagatcatcaattgtctgactcccgaaatgtctgc 1441
QY 1610 aagtgaagctttaccctgcagcagtagagtttagactaccagagagaatcttagagttcccg 1669
Db 1442 ccgtgaaaccccttccctgaaaaaccggacacgcccgcacaaagaagatttggaaatttccaa 1501
QY 1670 caagggatgtttatgttccaacaactacttacag 1703

Db 1502 cacgagaaagtatatatgtccctcacactgtgtacag 1535

Search completed: July 16, 2002, 12:24:50
Job time: 13897 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 16, 2002, 08:31:43 ; Search time 116.6 Seconds
(without alignments)
14384.096 Million cell updates/sec

Title: US-09-737-246-1
Perfect score: 6828
Sequence: 1 gtcgcgtcgcgcgcagcagc.....ctcgtcgcgcgcagcgctc 6828

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*

1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCFUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	64.2	0.9	7218	1 US-08-232-463-14	Sequence 14, Appl
C 2	45.2	0.7	7218	1 US-08-232-463-14	Sequence 14, Appl
C 3	44.2	0.6	2317	3 US-08-749-522-5	Sequence 5, Appl
C 4	44.2	0.6	3974	4 US-08-467-504-3	Sequence 3, Appl
C 5	41.6	0.6	1850	3 US-08-617-860B-32	Sequence 32, Appl
C 6	41.6	0.6	4098	2 US-08-605-106-4	Sequence 4, Appl
C 7	41.4	0.6	19124	2 US-08-487-826B-13	Sequence 13, Appl
C 8	40.6	0.6	6669	3 US-09-212-971-5	Sequence 5, Appl
C 9	40.6	0.6	6669	3 US-08-800-929A-5	Sequence 5, Appl
C 10	40.6	0.6	6669	4 US-09-617-053A-5	Sequence 5, Appl
C 11	39.6	0.6	860	4 US-08-998-416-287	Sequence 287, App
C 12	39.4	0.6	837	4 US-08-998-416-288	Sequence 288, App
C 13	39.4	0.6	1850	3 US-08-617-860B-32	Sequence 32, Appl
C 14	39.4	0.6	1939	1 US-07-715-751B-2	Sequence 2, Appl
C 15	39.4	0.6	4098	2 US-08-605-106-4	Sequence 4, Appl
C 16	39.2	0.6	1689	1 US-07-991-867B-41	Sequence 41, Appl
C 17	39.2	0.6	1689	2 US-08-544-332-41	Sequence 41, Appl
C 18	39.2	0.6	8457	1 US-07-991-867B-1	Sequence 1, Appl
C 19	39.2	0.6	8457	2 US-08-544-332-1	Sequence 1, Appl
C 20	39	0.6	2251	4 US-08-991-677-11	Sequence 11, Appl
C 21	38.8	0.6	289	4 US-09-007-005-17	Sequence 17, Appl
C 22	38.8	0.6	289	4 US-09-244-796-17	Sequence 17, Appl
C 23	38.8	0.6	665	2 US-08-883-795A-36	Sequence 36, Appl
C 24	38.8	0.6	5852	1 US-07-867-106-2	Sequence 2, Appl
C 25	38.6	0.6	658	4 US-08-998-416-595	Sequence 595, App
C 26	38.6	0.6	3047	1 US-07-927-851-2	Sequence 2, Appl
C 27	38.6	0.6	3047	1 US-08-453-323-2	Sequence 2, Appl

C 28	38.6	0.6	3047	1 US-08-440-520-2	Sequence 2, Appl
C 29	38.6	0.6	19124	2 US-08-487-826B-13	Sequence 13, Appl
C 30	38.4	0.6	289	4 US-09-007-005-17	Sequence 17, Appl
C 31	38.4	0.6	289	4 US-09-244-796-17	Sequence 17, Appl
C 32	38.2	0.6	662	4 US-08-936-165A-142	Sequence 142, App
C 33	38	0.6	2428	2 US-08-849-536A-3	Sequence 3, Appl
C 34	37.6	0.6	522	1 US-08-558-735-7	Sequence 7, Appl
C 35	37.6	0.6	522	4 US-08-906-480-7	Sequence 8, Appl
C 36	37.6	0.6	4612	2 US-08-447-031A-8	Sequence 1137, Ap
C 37	37.4	0.5	636	4 US-08-998-416-1137	Sequence 3, Appl
C 38	37.4	0.5	1107	1 US-08-165-038-3	Sequence 3, Appl
C 39	37.4	0.5	1107	1 US-08-621-081A-3	Sequence 3, Appl
C 40	37.4	0.5	1107	2 US-08-876-781-3	Sequence 3, Appl
C 41	37.4	0.5	1956	4 US-08-559-896B-1	Sequence 1, Appl
C 42	37.4	0.5	2110	4 US-09-419-459-1	Sequence 1, Appl
C 43	37.2	0.5	3618	1 US-07-872-678A-36	Sequence 36, Appl
C 44	37	0.5	2989	1 US-07-927-851-1	Sequence 1, Appl
C 45	37	0.5	2989	1 US-08-453-323-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgpt-F1s
; US-08-232-463-14

```
Query Match          0.9%; Score 64.2; DB 1; Length 7218;
Best Local Similarity 8.2%; Pred. No. 2.4e-07;
Matches 30; Conservative 197; Mismatches 140; Indels 0; Gaps 0;

QY 4132 gagtataagggaaaaagtggttgaacgaataagcttgaccttaagaatacaaa 4191
    :::::::::::::::::::: : :::::::::::::: : :::::::::::::: :
Db 1364 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1305

QY 4192 gacatgagagcaagcttgaagaagctatcttgygagcatagtgccaggaagaatg 4251
    :::::::::::::::::::: : :::::::::::::: : :::::::::::::: :
Db 1304 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1245

QY 4252 gtacggcgaagccgagacagctcgagagaagcccatctgyaagtgccttggaagtcaa 4311
    :::::::::::::::::::: : :::::::::::::: : :::::::::::::: :
Db 1244 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1185

QY 4312 gaaaattgaggtgagggaaagatatgactcactggtgcataaacaagagaagcttgac 4371
    :::::::::::::::::::: : :::::::::::::: : :::::::::::::: :
Db 1184 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1125

QY 4372 aaatcaagagcagagattgaacaggaagcactgattgataaaccctgctacagaagca 4431
    :::::::::::::::::::: : :::::::::::::: : :::::::::::::: :
Db 1124 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1065

QY 4432 aacctaatcatcttagatatacatagagattgtgtcagaccgtttcgttaacgaatcc 4491
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1064 CGCAAGCTCCCTCGACCTGCAGCCAAAGCTCGAATTATTTCTGTGAGCGTATGCAACG 1005

QY 4492 aaagaga 4498
    |||||
Db 1004 AAGGAAA 998

RESULT 2
US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
```

```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: PTZ9pt-F1s
US-08-232-463-14

Query Match          0.7%; Score 45.2; DB 1; Length 7218;
Best Local Similarity 6.8%; Pred. No. 0.048;
Matches 29; Conservative 211; Mismatches 184; Indels 0; Gaps 0;

QY 3202 cagaaggaacacagaatatgttgagagactcaataaagccttgcatctcttcaatgat 3261
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1012 CATACGCTCACAGAAATTATTCGAGCTTGCTGCGAGTCGAGGAGCTTGCGATYYYY 1071

QY 3262 ctgtgtctgttatgacagagattgtttttaagcctataaagtcctgctataaacag 3321
    :::::::::::::::::::: : :::::::::::::: : :::::::::::::: :
Db 1072 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1131

QY 3322 gtgtctcaagccttaccatcattaccgaatcccaagtgtctgtgtcttgaggctgat 3381
    :::::::::::::::::::: : :::::::::::::: : :::::::::::::: :
Db 1132 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1191

QY 3382 ttctacgaatcatctgcagtgcatgagcactatgtttacattaaacttaccctgcagctta 3441
    :::::::::::::::::::: : :::::::::::::: : :::::::::::::: :
Db 1192 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1251

QY 3442 cttaactcaactgcattccatcacctctgtttcttctgcaacatcgaagttctgga 3501
    :::::::::::::::::::: : :::::::::::::: : :::::::::::::: :
Db 1252 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1311

QY 3502 ttctcagaatgtlacaagaccaaagaatgtgcaaatatgttgaattacgltgccttc 3561
    :::::::::::::::::::: : :::::::::::::: : :::::::::::::: :
Db 1312 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1371

QY 3562 cggcaacagcattatltggcagagctgtgttaacagagctgctgtcatcttagacct 3621
    :::::::::::::::::::: : :::::::::::::: : :::::::::::::: :
Db 1372 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1431

QY 3622 gatg 3625
    :|
Db 1432 YYYG 1435

RESULT 3
US-08-749-522-5
; Sequence 5, Application US/08749522
; Patent No. 6096950
; GENERAL INFORMATION:
; APPLICANT: John, Maliyakal
; TITLE OF INVENTION: FIBER-SPECIFIC PROMOTERS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 411 East Wisconsin Avenue
; CITY: Milwaukee
; STATE: WI
; COUNTRY: U.S.A.
; ZIP: 53202-4497
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/749,522
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Baker, Jean C.
```



```

; REGISTRATION NUMBER: 35,433
; REFERENCE/DOCKET NUMBER: 670513.90244
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (414) 277-5709
; TELEFAX: (414) 271-3552
; INFORMATION FOR SEO ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2317 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-749-522-5

```

Query Match	0.68;	Score 44.2;	DB 3;	Length 2317;
Best Local Similarity	48.28;	Pred. No. 0.045;		
Matches 124;	Conservative	0;	Mismatches 133;	Indels 0;
				Gaps 0;

[illegible]

RESULT 4
US-08-467-504-3
; Sequence 3, Application US/08467504
; Patent No. 6211430
; GENERAL INFORMATION:
; APPLICANT: John, Maliyakal E.
; TITLE OF INVENTION: FBLATE PROMOTER
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 411 East Wisconsin Avenue
; CITY: Milwaukee
; STATE: WI
; COUNTRY: U.S.A.
; ZIP: 53202-4497
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467,504
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Baker, Jean C.
; REGISTRATION NUMBER: 35,433
; REFERENCE/DOCKET NUMBER: 110229.91152
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (414) 277-5709
; TELEFAX: (414) 271-3552
; INFORMATION FOR SEO ID NO: 3;
; SEQUENCE CHARACTERISTICS:
;

```

; LENGTH: 3974 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
;
US-08-467-504-3

```

Query Match	0.6%;	Score 44.2;	DB 4;	Length 3974;
Best Local Similarity	48.2%;	Pred. No. 0.063;		
Matches 124;	Conservative	0;	Mismatches 133;	Indels 0;
				Gaps 0;

QY	6537	caagaaataatc	aaatcgtagatt	gtttaa	cgcgttgaga	atcatg	cgctatg	gtttc	taatg	6596		
Db	1438	CAAAAAAAAA	ACTAATGTTG	GTGGTTGA	ATTTTATAT	TACGGA	TGTA	TGTATAT	TTT	1497		
QY	6597	tcggt	taacaagc	tgttat	ctttta	agacatt	ltaag	tactca	aaggt	tacata	cat	
Db	1498	TAA	ATAAAAT	TATGTTAT	TTAGATTCT	TATATAT	TTTGAG	CATTCC	ATACTAT	ATATCT	CG	
QY	6657	taccat	tattata	ccatag	ctaag	gtt	aaa	aattat	cactt	aaag	ttcgtat	ttt
Db	1558	TATACAT	ATATAT	TAAAA	ATAGTA	TATATA	AAAGTGA	TATTA	ACTTTAA	ATTACA	AGCATAA	
QY	6717	aatt	tata	tccac	cat	tata	tagatt	tcatt	tttg	gacc	cat	ttt
Db	1618	TAT	TAAAT	TTTGA	ATCA	TATTA	TTTTT	TATTTCT	ATTTAT	TTTAA	TTTAAT	TTT
QY	6777	ttaa	aggt	act	aaaaa							
Db	1678	TTCA	AAAT	TAA	AAAT	TTAA						

```

RESULT      5
US-08-617-860B-32/c
; Sequence 32, Application US/08617860B
; Patent No. 6133506
;
; GENERAL INFORMATION:
; APPLICANT: Tvpfer, R., Bautor, J., Bothmann, H., Fllsak,
; APPLICANT: Hvrlicke-Grandpierre, C., Klein, B., Martini, N
; APPLICANT: M ller, A., Schulte, W., Voeltz, M., Walek, J.,
; APPLICANT: Schell, J.
; TITLE OF INVENTION: Promoters
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Steinberg, Raskin & Davidson, P.C.
; STREET: 1140 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
;
; ZIP: 10036
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4 Mb storage
;
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/617,860B
; FILING DATE: 01-MAR-1996
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP94/02950
; FILING DATE: 05-SEP-1994
; APPLICATION NUMBER: DE P4329951.2
; FILING DATE: 04-SEP-1993
;
; INFORMATION FOR SEQ ID NO: 32:
;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1850 Base pairs
; TYPE: Nucleic acid
; STRANDEDNESS: Double stranded
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
;

```

```
ORIGINAL SOURCE:
ORGANISM: Cuphea lanceolata
IMMEDIATE SOURCE:
LIBRARY: genomic Lambda FIX II
CLONE: CITEg1
FEATURE:
NAME/KEY: CAAT-Signal
LOCATION: 1428..1432
FEATURE:
NAME/KEY: TATA-Signal
LOCATION: 1553..1556
FEATURE:
NAME/KEY: Transcription start
LOCATION: 1585
FEATURE:
NAME/KEY: Legumibox
LOCATION: 1642..1657
FEATURE:
NAME/KEY: Startcodon
LOCATION: 1797..1799
FEATURE:
NAME/KEY: CDS
LOCATION: 1797..1850
US-08-617-860B-32
```

```
Query Match      0.6%; Score 41.6; DB 3; Length 1850;
Best Local Similarity 48.7%; Pred. No. 0.21;
Matches 113; Conservative 0; Mismatches 119; Indels 0; Gaps 0;
```

```
QY 6542 aaatatacaaatgtagattgttaacgcttgagaatcatggtatggttctaatttcggg 6601
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 450 ATATATATAAATTCCTCCGGTATCTCCACAAAAATTTTTCATATTTCAAAAATTCAAA 391
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 6602 taacaagctgtatcttctaagacatttaactcaaaagtacactatacattacca 6661
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 390 AAATATTTAAAAACTTTTAAATATATTTTAAACCAAAAATAATTTTAAAAATTCACAA 331
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 6662 ttattataccatagctaaagtttaaaattatcacttaagttcgtatttttaatt 6721
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 330 ATATTTTAAAAATAAAAAAACTAATTTTAAATATTTTAAAAAATTTTAAAAAACTAATT 271
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 6722 ataccacattatagattcatttgaccatttaagttagtaatgctt 6773
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 270 TATAATATTTAAAAATATTTTAAAGAAATTTAAATATATTTAAATTT 219
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 6

```
US-08-605-106-4/C
; Sequence 4, Application US/08605106
; Patent No. 5910631
; GENERAL INFORMATION:
; APPLICANT: Topfer, R.
; APPLICANT: Martini, N.
; APPLICANT: Schell, J.
; TITLE OF INVENTION: MEDIUM CHAIN-SPECIFIC THIOESTERS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P.A.
; STREET: P.O. Box 2938
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/605,106
; FILING DATE: 23-SEPT-1996
; CLASSIFICATION: 800
```

```
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP94/02935
FILING DATE: 01-MAR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Woessner, Warren D
REGISTRATION NUMBER: 30,440
REFERENCE/DOCKET NUMBER: 235.001US1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 612-373-6900
TELEFAX: 612-339-3061
TELEX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 4098 Base pairs
TYPE: nucleic acid
STRANDEDNESS: double stranded
TOPOLOGY: linear
MOLECULE TYPE: : DNS (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Cuphea lanceolata
IMMEDIATE SOURCE:
LIBRARY: genomic Lambda FIX II
CLONE: CITEg1
FEATURE:
NAME/KEY: CDS
LOCATION: join(1797..2294, 2658..2791, 2898..3011, 3132
LOCATION: ..3303, 3391..3459, 3672..3941)
FEATURE:
NAME/KEY: Startcodon
LOCATION: 1797..1799
FEATURE:
NAME/KEY: exon II
LOCATION: 1787..2294
FEATURE:
NAME/KEY: Intron II
LOCATION: 2295..2657
FEATURE:
NAME/KEY: exon III
LOCATION: 2658..2791
FEATURE:
NAME/KEY: Intron III
LOCATION: 2792..2897
FEATURE:
NAME/KEY: exon IV
LOCATION: 2898..3011
FEATURE:
NAME/KEY: Intron IV
LOCATION: 3012..3131
FEATURE:
NAME/KEY: exon V
LOCATION: 3132..3303
FEATURE:
NAME/KEY: Intron V
LOCATION: 3304..3390
FEATURE:
NAME/KEY: exon VI
LOCATION: 3391..3459
FEATURE:
NAME/KEY: Intron VI
LOCATION: 3460..3671
FEATURE:
NAME/KEY: exon VII
LOCATION: 3672..3941
FEATURE:
NAME/KEY: Stopcodon
LOCATION: 3942..3944
US-08-605-106-4
```

```
Query Match      0.6%; Score 41.6; DB 2; Length 4098;
Best Local Similarity 48.7%; Pred. No. 0.34;
```

	Matches	113;	Conservative	0;	Mismatches	119;	Indels	0;	Gaps	0;
QY	6542	aaatatcaaatgtagatlgttaacgcttgagaatcatggtatggttctaatgttcggg	6601							
Db	450	ATATATATAAATTCCTCCGGTTATCTCCACACAAAATTTTTCATATTTCCAAAAATTCAAA	391							
QY	6602	taacaagctgtatcttttaagacattttaatgactcaaaagtacactatatacattacca	6661							
Db	390	AAATATTTAAAAACTTTTAATATATATTTTAAAAACCAAAAAATATTTTAAAAAATTCAACA	331							
QY	6662	ttatttataccatagctaaggttaaaaaatttattcactttaagtcgtattttaaatt	6721							
Db	330	ATATTTTAAAAATAAAAAAAACCTAATTTTAAATAATTTTAAAAAAATTTTAAAAAACTAATTT	271							
QY	6722	atatcaccatttatagatcatttggaccattttaaatgtagtaatgctt	6773							
Db	270	TATATATATTTAAAAATATATTTTTTAAAGAATTTAAATAATATTTTAAATTT	219							

```

RESULT 7
US-08-487-826B-13
; Sequence 13, Application US/08487826B
; Patent No. 5993827
; GENERAL INFORMATION:
; APPLICANT: Sim, Kim L.
; APPLICANT: Chitnis, Chetan
; APPLICANT: Miller, Louis H.
; APPLICANT: Peterson, David S.
; APPLICANT: Su, Xin-zhaun
; APPLICANT: Wellens, Thomas E.
; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
; TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbé Martens Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: California
; COUNTRY: US
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,826B
; FILING DATE: 10-SEP-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Israelsen, Ned
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: NIH121.001CPI
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19124 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
US-08-487-826B-13

Query Match 0.68; Score 41.4; DB 2; Length 19124;
Best Local Similarity 52.08; Pred. No. 1;
Matches 93; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

```

Db	15441	TTTATTAAAAATGAAAAAAGAAAAATGAAATATATAAAAAAAATTTATTAATAATAAAAAAA	15500
QY	377	gtatacagaagactggcaatgtgtcatcagaatatatcataaatctgggaacagattta	436
Db	15501	AAAAAAAAAAAAAGAGAAAAATTTTTTTAAAAAATAATATAAAAAATTTATAATAAATAATATAA	15560
QY	437	atcccaatacatctagataaacagaaagaaagccaataagtgctggccaataacaagttlct	495
Db	15561	ATTTTGATAGAAATAAAAAATGAAMMAGATTCATAAAAAAAATTTAAAAAATAATTTTAT	15619

```

RESULT      8
US-09-212-971-5/c
; Sequence 5, Application US/09212971B
; Patent No. 6107041
;
; GENERAL INFORMATION:
; APPLICANT: Korneluk, Robert G
; APPLICANT: Mackenzie, Alexander E
; APPLICANT: Liston, Peter
; APPLICANT: Baird, Stephen
; APPLICANT: Tsang, Benjamin K
; APPLICANT: Pratt, Christine
; TITLE OF INVENTION: DETECTION AND MODULATION OF IAPS AND
; TITLE OF INVENTION: NAIP FOR THE DIAGNOSIS AND TREATMENT OF PROLIFERATIVE
; TITLE OF INVENTION: DISEASE
; FILE REFERENCE: 07891/009002
; CURRENT APPLICATION NUMBER: US/09/212,971B
; CURRENT FILING DATE: 1998-12-16
; EARLIER APPLICATION NUMBER: 60/017,354
; EARLIER FILING DATE: 1996-04-26
; EARLIER APPLICATION NUMBER: 60/030,590
; EARLIER FILING DATE: 1996-11-14
; EARLIER APPLICATION NUMBER: 08/800,929
; EARLIER FILING DATE: 1997-02-13
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 6669
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: variation
; LOCATION: (3677)..(3951)
; OTHER INFORMATION: n can be any nucleotide
US-09-212-971-5

```

Query Match	0.68;	Score 40.6;	DB 3;	Length 6669;
Best Local Similarity	49.8%;	Pred. No. 0.87;		
Matches 103;	Conservative 0;	Mismatches 104;	Indels 0;	Gaps 0;
QY 6575	tcacgtcatggttctcaatgltcgggtaacaagctgttatcctttaagacattttaatg	6634		
DB 2425	TCCAATGCTTTGTTTCCAGGTGGCAGAGAAACATCACATCTCAATTACCAACAATTATTTC	2366		
QY 6635	actcaagggtacacatatacatttaaccatttataaccatagctaaaggtttaaatttat	6694		
DB 2365	CCCTCAAGGTTAAGTAATTGTATCTAACTTCATATATTCCTTCCCTGGCCAAAAATTTCAC	2306		
QY 6695	tcacatttaagttcgtatctttttaattatatacaacatttatagatttcatttgyaaccat	6754		
DB 2305	ACATTTGTTATTAATAAATCTTCTTTGTTATTCATCACCTGCTTTATATTAATTTCTTAAGTTT	2246		
QY 6755	tttaaatgtagtaatgcttattttaaa	6781		
DB 2245	CTTCATTTCTTAATAATGTAACCTTCAA	2219		


```

;
; COMPUTER READABLE FORM:
;
; MEDIUM TYPE: Floppy disk
;
; COMPUTER: IBM PC compatible
;
; OPERATING SYSTEM: PC-DOS/MS-DOS
;
; SOFTWARE: Patentin Release #1.0, Version #1.30
;
; CURRENT APPLICATION DATA:
;
; APPLICATION NUMBER: US/08/998,416
;
; FILING DATE: 24-DEC-1997
;
; CLASSIFICATION: 435
;
; PRIOR APPLICATION DATA:
;
; APPLICATION NUMBER: CH 0016/97
;
; FILING DATE: 31-DEC-1996
;
; ATTORNEY/AGENT INFORMATION:
;
; NAME: Meigs, J. Timothy
;
; REGISTRATION NUMBER: 38,241
;
; REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
;
; TELECOMMUNICATION INFORMATION:
;
; TELEPHONE: 919-541-8587
;
; TELEFAX: 919-541-8689
;
; INFORMATION FOR SEQ ID NO: 287:
;
; SEQUENCE CHARACTERISTICS:
;
; LENGTH: 860 base pairs
;
; TYPE: nucleic acid
;
; STRANDEDNESS: single
;
; TOPOLOGY: linear
;
; MOLECULE TYPE: DNA (genomic)
;
; ORIGINAL SOURCE:
;
; ORGANISM: PAG1240UP
;
; US-08-998-416-287
```

```

Query Match          0.6%; Score 39.6; DB 4; Length 860;
Best Local Similarity 56.0%; Pred. No. 0.46;
Matches 75; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

QY 6662 ttattataccatagctaaggttaaaatttattcactttaagtcgtatttttaattt 6721
    |||| ||| | | | | |||| | | | | |||| | | | | ||||
Db 443 TTTATTAAATAAAAATGTAATTTTAATATTTAAAAATTAAATTAATACTATTTAATAATAT 502

QY 6722 atatcaccatttatagattcaatttggaaccattttaaatgtagtaatgcttatttaa 6781
    ||| | | | | | | | | |||| | | | | |||| | | | | |||
Db 503 CTATAGTAATTTCTTATTATTATTATAACATTTTAAATGTTTATGTTTAATAGATA 562

QY 6782 ggtactaaaaata 6795
    || | |||||
Db 563 ATACCAATTAAATA 576
```

```

RESULT 12
US-08-998-416-288
; Sequence 288, Application US/08998416
; Patent No. 6239264
; GENERAL INFORMATION:
; APPLICANT: Philippsen, Peter
; APPLICANT: Pohlmann, Rainer
; APPLICANT: Steiner, Sabine
; APPLICANT: Mohr, Christine
; APPLICANT: Wendland, Jurgen
; APPLICANT: Knechtle, Philipp
; APPLICANT: Rebischung, Corinne
; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPII
; TITLE OF INVENTION: AND USES THEREOF
; NUMBER OF SEQUENCES: 1152
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6239264artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: No. 6239264th Carolina
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

```

;
; OPERATING SYSTEM: PC-DOS/MS-DOS
;
; SOFTWARE: Patentin Release #1.0, Version #1.30
;
; CURRENT APPLICATION DATA:
;
; APPLICATION NUMBER: US/08/998,416
;
; FILING DATE: 24-DEC-1997
;
; CLASSIFICATION: 435
;
; PRIOR APPLICATION DATA:
;
; APPLICATION NUMBER: CH 0016/97
;
; FILING DATE: 31-DEC-1996
;
; ATTORNEY/AGENT INFORMATION:
;
; NAME: Meigs, J. Timothy
;
; REGISTRATION NUMBER: 38,241
;
; REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
;
; TELECOMMUNICATION INFORMATION:
;
; TELEPHONE: 919-541-8587
;
; TELEFAX: 919-541-8689
;
; INFORMATION FOR SEQ ID NO: 288:
;
; SEQUENCE CHARACTERISTICS:
;
; LENGTH: 837 base pairs
;
; TYPE: nucleic acid
;
; STRANDEDNESS: single
;
; TOPOLOGY: linear
;
; MOLECULE TYPE: DNA (genomic)
;
; ORIGINAL SOURCE:
;
; ORGANISM: PAG1241RP
;
; US-08-998-416-288
```

```

Query Match          0.6%; Score 39.4; DB 4; Length 837;
Best Local Similarity 43.9%; Pred. No. 0.52;
Matches 169; Conservative 0; Mismatches 216; Indels 0; Gaps 0;

QY 6396 aaatgcatcctctaaactgaatgcactgttttatctcatcgcgaagaagccatgtatca 6455
    |||| | | | | | | | | | | | | | | | | | | | | | |
Db 187 AAAAATTAATTTTAATATGAAATACATTTTAGTCTATGTTCAAAATTTTAAATAGTTATTA 246

QY 6456 acatcgatgctgaaaagatctatttggaaaccaacatggaatggaatctctggaattat 6515
    | | | | | | | | | | | | | | | | | | | | | | | |
Db 247 AAATATTATAGATTTATTATTCTTTTAAATAATTAATAGATTATCAATTAATTA 306

QY 6516 ttcatlgaagaatgcagtgccgaagaaatatcaaatgtagaattgttaacgcttgagaat 6575
    | | | | | | | | | | | | | | | | | | | | | | | |
Db 307 ATATATTATTATTAAATGTTTATTAAATAATATATATTATTATTAATAAGATTAAATT 366

QY 6576 catgctatggttctaatgttcggytaacaagcgttatcttctaagacatttaatga 6635
    || | | | | | | | | | | | | | | | | | | | | | |
Db 367 TATTTAATATTTGTAATTTATTATTATTATTAATATATATCTATTTTTAATATATTGTT 426

QY 6636 ctcaaagttacatatcatttaccatttatataccatagctaaaggttaaaatttaatt 6695
    | | | | | | | | | | | | | | | | | | | | | | | |
Db 427 GATTATATTAATTAATCTTTTATTAAGCAATTATTATTAATAAATTAATTTTAACCTTAATT 486

QY 6696 cactttaagttcgtattttttaatttatataccatttatagattcatttggaccatt 6755
    | | | | | | | | | | | | | | | | | | | | | | | |
Db 487 TCTTATTATTATTATTATATTATTATTAATAAATTATATTTCATTATTATTATTATTAT 546

QY 6756 ttaatgtagtaatgcttatttaa 6780
    |||| | | | | | | | | | | | | | | | | | | | | | |
Db 547 TTAATTAATTAATTAATTATTATTAATTA 571
```

```

RESULT 13
US-08-617-860B-32
; Sequence 32, Application US/08617860B
; Patent No. 6133506
; GENERAL INFORMATION:
; APPLICANT: Tvpfer, R., Bautor, J., Bothmann, H., Filsak, E.,
; APPLICANT: Hviricke-Grandpierre, C., Klein, B., Martini, N.,
; APPLICANT: M iller, A., Schulte, W., Voeltz, M., Walek, J.,
; APPLICANT: Scheil, J.
; TITLE OF INVENTION: Promoters
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
```

ADDRESSEE: Steinberg, Raskin & Davidson, P.C.
STREET: 1140 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4 Mb storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/617,860B
FILING DATE: 01-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP94/02950
FILING DATE: 05-SEP-1994
APPLICATION NUMBER: DE P4329951.2
FILING DATE: 04-SEP-1993
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 1850 Base pairs
TYPE: Nucleic acid
STRANDEDNESS: Double stranded
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Cuphea lanceolata
IMMEDIATE SOURCE:
LIBRARY: genomic lambda FIX II
CLONE: CITEg1
FEATURE:
NAME/KEY: CAAT-Signal
LOCATION: 1428..1432
FEATURE:
NAME/KEY: TATA-Signal
LOCATION: 1553..1556
FEATURE:
NAME/KEY: Transcription start
LOCATION: 1585
FEATURE:
NAME/KEY: Legumibox
LOCATION: 1642..1657
FEATURE:
NAME/KEY: Startcodon
LOCATION: 1797..1799
FEATURE:
NAME/KEY: CDS
LOCATION: 1797..1850
US-08-617-860B-32

Query Match 0.6%; Score 39.4; DB 3; Length 1850;
Best Local Similarity 50.3%; Pred. No. 0.85;
Matches 97; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

QY 6614 atccttaagacattttaaagctcaaggtacactatacatttaccattattatacca 6673
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 209 ATATTTTAAATTTAAATATATTTTAAATTCCTTTAAAAAAATATTTTAAATATTA 268

QY 6674 tagctaaggttaaaaatttaactcatttaagtcglatctttaaattatataccacattt 6733
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 269 TAAAAATTAGTTTTTAAAAATTTTAAATATTTTAAATTTAGTTTTTTTAAATTTTAA 328

QY 6734 atagattcatttttgaccacattttaaagttagtaatgctattttaaaggtactaaaaa 6793
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 329 ATTGTTGAATTTTAAAAATTTTGTGTTTAAAAAATATATTTTAAAGTTTTTAAATAT 388

QY 6794 taatgtaatgttt 6806
|| ||||| |||||
Db 389 TTTTGAATTTT 401

RESULT 14
US-07-715-751B-2/c
Sequence 2, Application US/07715751B
Patent No. 5391725
GENERAL INFORMATION:
APPLICANT: CORUZZI, GLORIA M
APPLICANT: EDWARDS, JANICE W
APPLICANT: WALKER, ELSEBETH L
APPLICANT: BREARS, TIMOTHY B
TITLE OF INVENTION: NOVEL ORGAN-SPECIFIC PLANT PROMOTER
TITLE OF INVENTION: SEQUENCES
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: PENNIE & EDMONDS
STREET: 1155 AVENUE OF THE AMERICAS
CITY: NEW YORK
STATE: N.Y.
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/715,751B
FILING DATE: 19910613
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: MISROCK, S. LESLIE
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 3288-017-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1939 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-07-715-751B-2

Query Match 0.6%; Score 39.4; DB 1; Length 1939;
Best Local Similarity 46.8%; Pred. No. 0.87;
Matches 124; Conservative 0; Mismatches 141; Indels 0; Gaps 0;

QY 6540 gaaatatcaaatgtagattgttaacgcttgagaatcatgctatglttctaagtctg 6599
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1212 GAATTTCTCTATTATAGATGACTTAATTAAGCAATTAAGTTTATGTTAGTATCTTAT 1153

QY 6600 ggtacaagctgttactctttaaagacatttaagtactcaaggtacacatacatttac 6659
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1152 AGTTAGTTTATTTATTTATATATGAATTAAGTATTAAGTTAAGATCATTTATTTATTT 1093

QY 6660 cattattataccatagctaaggttaaaaaattatctcatttaagttcgtattttaat 6719
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1092 TAGTTTATTTCATGTATTTAATAAAGAAATTTAATTTGAATAATAGTGTCTCTATTTTAT 1033

QY 6720 ttatatcaccatttataagattcattttggaccacattttaaagttagtaatgctattt 6779
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1032 TTGATTTTTCATTTTGTGATTTTATTTTAAATTCAAATTAATTTTCTTTTCTTTTGA 973

QY 6780 aaggtactaaaaaataatgtgaatgt 6804
|| ||||| ||||| |||||
Db 972 ACTTGAAATATTTTCTTTGTACAACT 948

```

RESULT 15
US-08-605-106-4
; Sequence 4, Application US/08605106
; Patent No. 5910631
; GENERAL INFORMATION:
; APPLICANT: Topfer, R.
; APPLICANT: Martini, N.
; APPLICANT: Schell, J.
; TITLE OF INVENTION: MEDIUM CHAIN-SPECIFIC THIOESTERS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schweigman, Lundberg, Woessner & Kluth, P.A.
; STREET: P.O. Box 2938
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/605,106
; FILING DATE: 23-SEPT-1996
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP94/02935
; FILING DATE: 01-MAR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Woessner, Warren D
; REGISTRATION NUMBER: 30,440
; REFERENCE/DOCKET NUMBER: 235.001US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-373-6900
; TELEFAX: 612-339-3061
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4098 Base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double stranded
; TOPOLOGY: linear
; MOLECULE TYPE: : DNS (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Cuphea lanceolata
; IMMEDIATE SOURCE:
; LIBRARY: genomic lambda FIX II
; CLONE: ClTeg1
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(1797..2294, 2658..2791, 2898..3011, 3132
; LOCATION: ..3303, 3391..3459, 3672..3941)
; FEATURE:
; NAME/KEY: Startcodon
; LOCATION: 1797..1799
; FEATURE:
; NAME/KEY: exon II
; LOCATION: 1787..2294
; FEATURE:
; NAME/KEY: intron II
; LOCATION: 2295..2657
; FEATURE:
; NAME/KEY: exon III
; LOCATION: 2658..2791
; FEATURE:
; NAME/KEY: intron III
; LOCATION: 2792..2897
; FEATURE:
; NAME/KEY: exon IV
; LOCATION: 2898..3011

```

```

; FEATURE:
; NAME/KEY: intron IV
; LOCATION: 3012..3131
; FEATURE:
; NAME/KEY: exon V
; LOCATION: 3132..3303
; FEATURE:
; NAME/KEY: intron V
; LOCATION: 3304..3390
; FEATURE:
; NAME/KEY: exon VI
; LOCATION: 3391..3459
; FEATURE:
; NAME/KEY: intron VI
; LOCATION: 3460..3671
; FEATURE:
; NAME/KEY: exon VII
; LOCATION: 3672..3941
; FEATURE:
; NAME/KEY: Stopcodon
; LOCATION: 3942..3944
US-08-605-106-4

```

Query Match	0.6%;	Score 39.4;	DB 2;	Length 4098;
Best Local Similarity	50.3%;	Pred. No. 1.4;		
Matches 97;	Conservative 0;	Mismatches 96;	Indels 0;	Gaps 0;

QY	6614	atcctttaagacattttaaatagcactcaaaagylacactataacattaccattattaccca	6673
Db	209	ATAATTTTAAAAATTTAAATATATTTTAAATCTTTAAAAAAATATTTTAAATATTA	268
QY	6674	tagctaaggttaaaaatttatcacttaagttcgatttllttaattatataccacatt	6733
Db	269	TAAAAATTAGTTTTTAAAAATTTTTTAAATATTTTAAAAATTAGTTTTTTTATTTTAAAAAT	328
QY	6734	atagattcatttlttgaccacattttaaatytagtaatgcttattttaaagytaactaaaaa	6793
Db	329	ATTGTTGAAATTTTAAAAATATTTTTTGCTTTTAAAAATATATTTTAAAGCTTTTAAATAT	388
QY	6794	tatgtgaatgctt	6806
Db	389	TTTTTGAATTTT	401

Search completed: July 16, 2002, 12:12:08
Job time: 13225 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 16, 2002, 08:30:28 ; Search time 7983.09 Seconds
(without alignments)
17898.632 Million cell updates/sec

Title: US-09-737-246-1

Perfect score: 6828
Sequence: 1 gtcgcgcgtcgcgcgcagcagc.....ctcgtgcgcgcagggcctc 6828

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htgo_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
------------	-------	-------------	--------	-------	-------------

1	6828	100.0	6828	6	AX172874	AX172874 Sequence
2	4224	61.9	4311	9	AB051558	AB051558 Homo sapi
3	4123	60.4	4144	6	AX172880	AX172880 Sequence
4	4123	60.4	4144	6	AX172932	AX172932 Sequence
5	4123	60.4	4200	6	AX172980	AX172980 Sequence
6	3230.8	47.3	3253	6	AX172978	AX172978 Sequence
7	3201.8	46.9	3327	9	AK055905	AK055905 Homo sapi
8	2360.2	34.6	2452	9	AK055493	AK055493 Homo sapi
9	2132.8	31.2	6372	6	AX173175	AX173175 Sequence
10	1821	26.7	2607	9	AK055401	AK055401 Homo sapi
11	1749	25.6	7215	6	AX174569	AX174569 Sequence
12	1552.8	22.7	4886	9	AB037816	AB037816 Homo sapi
13	1223.6	17.9	4577	9	AK024436	AK024436 Homo sapi
14	1119.4	16.4	2825	9	BC008335	BC008335 Homo sapi
15	1079	15.8	4026	6	AX174571	AX174571 Sequence
16	1079	15.8	4027	6	AX174664	AX174664 Sequence
17	919.6	13.5	2149	6	AX173233	AX173233 Sequence
18	853.4	12.5	1900	10	BC021414	BC021414 Mus muscu
19	843.2	12.3	2014	6	AX174662	AX174662 Sequence
20	791	11.6	2904	9	BC019102	BC019102 Homo sapi
21	622.6	9.1	73515	2	AC017375	AC017375 Drosophil
22	621.6	9.1	166626	3	AC008318	AC008318 Drosophil
23	621.6	9.1	301639	3	AE003590	AE003590 Drosophil
24	546.6	8.0	2219	9	AK055296	AK055296 Homo sapi
25	436.6	6.4	1242	9	AF194407	AF194407 Homo sapi
26	433.2	6.3	100190	2	AL391257	AL391257 Homo sapi
27	433.2	6.3	127370	9	AL451044	AL451044 Human DNA
28	433.2	6.3	131617	2	AC096946	AC096946 Homo sapi
29	407	6.0	473	6	AX341253	AX341253 Sequence
30	363.2	5.3	382	6	AX341959	AX341959 Sequence
31	334.4	4.9	933	6	AX172954	AX172954 Sequence
32	324.6	4.8	1566	9	AK022097	AK022097 Homo sapi
33	308	4.5	3152	9	AK056684	AK056684 Homo sapi
34	305.4	4.5	2299	9	AK054649	AK054649 Homo sapi
35	304.8	4.5	2610	6	AX058192	AX058192 Sequence
36	304.8	4.5	6454	6	AX173022	AX173022 Sequence
37	301.6	4.4	3472	6	AX173120	AX173120 Sequence
38	298.4	4.4	4391	6	AX173028	AX173028 Sequence
39	298.4	4.4	4393	6	AX173118	AX173118 Sequence
40	291	4.3	139887	9	CNS01RGX	AL160233 Human chr
41	244.4	3.6	100190	2	AL391257	AL391257 Homo sapi
42	242.4	3.6	185624	9	AC103923	AC103923 Homo sapi
43	230.6	3.4	1669	9	HSM802936	AL583913 Homo sapi
44	216.2	3.2	260	6	AX172960	AX172960 Sequence
45	214.6	3.1	3899	6	AX058220	AX058220 Sequence

ALIGNMENTS

RESULT 1
AX172874 LOCUS AX172874 6828 bp DNA linear PAT 03-JUL-2001
DEFINITION Sequence 1 from Patent WO0142297.
ACCESSION AX172874
VERSION AX172874.1 GI:14597911
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 6828)
AUTHORS Lu,P., Garman,J.D. and Candia,A.F.
TITLE Clasp-3 transmembrane protein
JOURNAL Patent: WO 0142297-A 1 14-JUN-2001;
Arbor Vita Corporation (US)
FEATURES
source location/Qualifiers
1..6828
/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="full-length human CLASP-3 CDNA"
22..6294
/note="human CLASP-3"

/codon_start-1
 /protein_id="CAC43726.1"
 /db_xref="GI:14597912"
 /translation="MAERRAFAQKISRVAAEVRKQISGQSGSPQLKNIIVGNISSHTHTVLTLEAVDPVDLEDYLITHPLEAVSGPLRLDIEFPDDIEVVVSPDRCRTLVA
 VPEESMDPHVRCIRSXTEDWALVIRKHYKLGTFGNPTLDKQERQGLPKQFESS
 DEAPDNGNSQDDQDKELFRHSMISIDTPKRSWASIFDLKNSIPDALNPLDRTNEE
 IDRONDOQRKSNRHKDLFALHRSMPDEEDETLERSVPDKPHEFGRLPLVCLSKFEI
 EIEPIASLALYDVKEKKISENFYFDLNSQMKGLRPHVPAAITTLARSAIFSTI
 YPSQDAVLVIKLEKVLQOGDIGECAEPYIMFEADATKNKELEKLSAQDOFCQRLG
 KYRMPACNTLSERPATLPTLVNFQKQEGRLDSDDEKFLADMRPSPVLRRLRPTAQ
 TSGDADNLTSEFPATLTPVLEKLVDPDSRVLPTRILEFPARDVVPNTYRNLLYI
 LKIDISAPENPHYCLTPELLQVKLXDPDSRVLPTRILEFPARDVVPNTYRNLLYI
 YPSQLNEANRQGSARNITVKQFMGDEPSNAMPVIFGKSSCSFEKAVTAVVYHNR
 SPDEKLEKVLKLPATLTDHHLLETFYHVCQDKONTPLETPGYTWTIMLQNGRLKT
 GQCFDPLVLEKLPQAYSLSPEVLPBGMKVVDHKKQFNVEVAVSSITHQDPLDKF
 FALVALDEHLEFPVRIGDMRIMENNELMELKSSISALNSQLPEPVRLDLDDKLLI
 LVIRPEVIAQIVNLGQASFEAMASTINRLKHNLEGNHDQHRNSLASYIHVFRLP
 NITPNSSSPGPGGLGSGVHYATMARSAVNPASLNRSRLSNSNPDISGTPSPDE
 VRSILGSKGLDRSNSVWNTGQCPKAPAWGNSNPSPAESTQAMDRSCNMSHTETSLE
 QTLGRLPTKLEHEELALQWVSGSVRESALQQAWEFFELMWKSVYHHLFYNDKLE
 APRKSRPERPEMDIALVSTIASDIAVSRFQKDEMERLNTSLAFPLNDLVSMDRG
 FVPSILKSCYKQOVSSKLYSLPNPSVYSLRLDLFRLICSHENYVTLNPLCSLLTPPAS
 PPSVSSATSQSGESTNWQDQKIANPPELFRQOHLYAGLIVLELAVILDDPAEG
 LFGHKKVINMVHNLSSHSDSPRSDPQIKARVAMLYPLIGLIMETVPLQDYFET
 HNOGRPCIAATDYESGSMISOVAMAIAGTSVLPOLTRPGSFLLTSGROHTEF
 SAERSRLILCLLWLNKNADEYLOKWFDTLSVLQNLRLDLLYLCVSCFEYKGVKVF
 ERKMSLTFFKSKMDRAKLEALIGSIGARQEMVRRSGLERSPSSAFSGQENLWR
 KDMTHWRQNTSEKDKSRAIEHEHALDGNLATEANLILDTLEIYQVVSYESKESI
 LGLEVLKVLHSMACNQSAYVLOHCFATQALVSKFPELLFEETEOCADCLRLRHG
 SSSIGTIRSHPSASLYLMRQNEFEGNNFARVKMQVPMSSLVGTSONFNEEFLRS
 KTLITLYAEDLELRETFPDQOVDLNLHMLISTVTKMKEHOEDEMLIDLMYRIA
 KQVOTSPERLTWLONMAGKHSESNHAAEQCLVHSAALVAEYLSMLEDRKYLPVGC
 TFQNISSNLVEESAVSDVVSDEEGICSGKYFTESGLVGLLEQAAESFMAGYEA
 NEYVKVLIPIHEANRDAKLSITHGKLQEAFSKIVHOSGTGWRNFGTYFRVGEYGTKE
 GDLDQEEFVYKEPAITKLAIESIRHLEGFYGERGEDVEVYKISNVDKCLDKPNKAY
 IQITVEYFPDYTEMKDRIITYFDKNYNLRFMYCTPTFLDGRAHGEHLHQFKRKTILT
 TSHAFFYIKTRVNTTHKEEIIITPIEVAIEDMOKTQELAPATHODPADPKMLQVLO
 GSVGTIVNQGPLEVAQVFLSELPSPDKLFRHHNKLRLCLCFKIDFTKCEDALRKNSLIG
 PVQKEYQRELKLSSP"
 1384 C 1432 G 1930 T

	Query Match	Best Local Similarity	Score 6828;	DB 6;	Length 6828;	
	Matches 6828;	Conservative	0;	Mismatches	0;	Indels
						Gaps
QY	1	gtccgcgctgcgccgacgagccaatgyccgagcgccgcgcttcggcccagaatacagcaga	60			
Db	1	GTGCGCCGTGCGCCGCAGCAGCCCATGGCCGAGCGCGGCCCTTCGCCACAGAATCAGCAGA	60			
QY	61	acggtgycagccgaagttagaaagcagatctccygacaatatagtglttcccccaactg	120			
Db	61	ACGGTGGCAGCCGAAGTTAGGAAGCAGATCTCCGGACAATAATAGTGTTCTCCCCAACTG	120			
QY	121	ctcaaaaaaccttaatatltgttygcaatatatcccatcacaccacagtgccccctaaccga	180			
Db	121	CTCAAAAA CCTTAATATTGTGTGGCAATATATCCCATCACACCACAGTGCCCTTACCGAA	180			
QY	181	gcagtagatccagatgtgatatttgygaagattaccctcatctaccatcctttgctgtgatltct	240			
Db	181	GCAGTAGATCCAGTGGATTTGGAAGATTACCTCATTTACTTCATCCTTTGGCTGTGGATTC	240			
QY	241	gggccttacgcygatttgttatgtaatlctccaccagatgatatgtgaagttgtttatatgtcct	300			
Db	241	GGGCTTTACGGGATTTGATTGAATTTCCCTCCAGATGATATTGAAGTTGTTTATAGTCT	300			
QY	301	cgggactgcagaactctctgttttcagctgtacctggaagaagtgaaatgatacaccatgtt	360			
Db	301	CGGGACTGCAGAACTCTTGTTTCAGCTGTACCTGAAGAAAGTGAATGGATCCACATGTT	360			
QY	361	agagactgtataagaagtatacagaagaactggygcaattgttcatalcagaanaatatcataaa	420			

Db	361	AGAGACTGTATAAGAGTTATACAGAGACTGGGCAATTGTCATCAGAAAAATATCATAAA	420
QY	421	ttgggaacagagattcaatcccaatatacatagataaacaagaagaagcacaagaagtgtg	480
Db	421	TTGGGAACAGGATTTAATCCCAATACATTAGATAAACAAGAAAGCAAAAAGGTTTG	480
QY	481	ccaacaagafttttgaatctgatgaagctccagatgycacaagctaccagatgatcaa	540
Db	481	CCAAACAAGTTTTGAAFTCTGATGAAGCTCCAGATGGCAACAGCTACCAGATGATCAA	540
QY	541	gatgaccttaaaagacgttcaatgtcaatagatagatacccaagggtagctggcctgt	600
Db	541	GATGACCTTAAGAAGACGTTCATGTCAATAGATGATACCCCAAGGGGTAGCTGGCCTGT	600
QY	601	agtatcttgacttgaaaaaattcaacttcctgatgcttggcttcccaatttacttgatga	660
Db	601	AGTATCTTGGACTTGAAAAATTCACTTCCTGATGCTTTGGCTTCCCAATTTACTTGATCGA	660
QY	661	actccaatgaaagaaatagaccgltcagaatgatgaccaaagaaatcaaacgcgtacaaa	720
Db	661	ACTCCAAATGAAGAATAAGACCGTCAGAATGATGACCAAGAATCAAACCGTCACAAA	720
QY	721	gaacttttgccttgcataccatcacaatgaggaagaaccaatagaaacgcgttagtgtt	780
Db	721	GAACCTTTTGCCTTGGCATCCATCACCAGATGAGGAAGAACCAATAGAACGGCTTAGTGTT	780
QY	781	cctgatatacccaagaacaatttgytcaaaagactcttglyaaaatgtctatcaactcaag	840
Db	781	CCTGATATACCCCAAGAACATTTTGGTCAAAAGACTTCTTGTAAMATGCTTATCACTCAAG	840
QY	841	tttgaaaattgaaaaattgaaccatttltgcaaglttggccttatatgatgtcaaggaag	900
Db	841	TTTGAATTTGAATTGAACCCATTTTTGCAAGTTTGCTTTATATGATGTCAAGCAAAAAG	900
QY	901	aaaaagattcagaaaaacttlattttgaccttaattctgacagatgaagggltgtta	960
Db	901	AAAAAGATTTCAGAAAACTTTTATTTTGACCTTAATTCTGAGCAGATGAAGGGTGTTA	960
QY	961	cytccacatgtaccacactgctgccattactaccctgygcaagatcagcaatttttctatc	1020
Db	961	CGTCCACATGTACCACTGCTGCCATTACTACCCCTGGCAAGATCAGCAATTTTCTCATC	1020
QY	1021	acttatccttcccaagatgttttctctgtaataaagctagaaaaagtcctacagcaagga	1080
Db	1021	ACTTATCCTTCCCAAGATGTTTTCTTTGTAATAAAGCTAGAAAAAGTCTTACAGCAAGGA	1080
QY	1081	gacattgagagagtgctgcagaaccatatatgtatttcaaaagaagcagatgccaaccaagaat	1140
Db	1081	GACATTGGAGAGTGTGCAGAACCATATATGATTTTCAAAAGAAGCAGATGCCCAAGAAT	1140
QY	1141	aaagaaaaaactggaagaaactgaagaagtcgaagcagatcagtttgcgaagacttgggaaa	1200
Db	1141	AAAGAAAAAAGTGGAGAAACTGAAGAGTCAAGCAGATCAGTTTGGCCAAGACTTGGGAAA	1200
QY	1201	tatcgcatgccttttgccttgagctgcaatccattlaatgaatatgtttagcagtgctggg	1260
Db	1201	TATCGCATGCCTTTTGGCTTGAGCTGCATCCATTATGAATATTTGTTAGCAGTGTGGG	1260
QY	1261	agtttggaaagagatctctacagaagtagaatacagtaactgagagaacgaaaaagggcttgg	1320
Db	1261	AGTTTGGAAAGAGATTCTACAGAAGTAGAAATCAGTACTGGAGAACGAAAAAGGGCTTGG	1320
QY	1321	tcagaagaggaggaattctagatattgtgycagaacgatcacttgaagaagacaacaagtga	1380
Db	1321	TCAGAGAGGAGGAATTCTAGTATTTGTGGCAGACGATCAGTTGAAAAGACAACAAGTGGA	1380
QY	1381	gatgatgcttglytaacttgacagacttgcaccagctactctcacagtgacaattttttt	1440
Db	1381	GATGATGCTTGTAACTTGACGAGCCTTTCGACCAGCTACTCTCACAGTGACAAATTTT	1440
QY	1441	aagcaggaagagaccgcttaagtgtatgaagatctctacaatttccttgcgtatatgaga	1500

Db 1441 AAGCAGGAGAGAGACCGCTTAAGTGAAGANCTCTACAATTCCTTGCTGATATGACA 1500
QY 1501 aggcacatctctgtcttaacggcgactaagaacctattacagctcagctcaagatagacatt 1560
Db 1501 AGGCCATCTTCTGTCTTACGGCGACTAAGACCTATTACAGCTCAGCTCAAGATAGACATT 1560
QY 1561 tctcccgcaacctgnaaatcccccattattgctctaactccggagctgcttcaagtgaagctt 1620
Db 1561 TCTCCCGCACCTGAAAATCCCATTTATTGCTTACTCCGGAGCTGCTTCAAGTGAAGCTT 1620
QY 1621 taacctgacagtagagttagaacctaccagagaaatcttagagttcccgcaaggagatt 1680
Db 1621 TACCCTGACAGTAGAGTTAGACCTTACCAGAGAAATCTTAGAGTTTCCCGCAAGGAGTGT 1680
QY 1681 tatgttccaaacaactacttacagaaaatctctctacataataccctcagagtttaattt 1740
Db 1681 TATGTTCCAAAACACTACTTACAGAAATCTTCTACATATACCCTCAGAGTCTTAATTTT 1740
QY 1741 gccaatcgltcaaggttcgtctagaanaataaacagtgaaagtcagtttatgtatggagag 1800
Db 1741 GCCAATCGTCAAGGTTCTGTAGAAATATTAACAGTGAAGTCCAGTTTATGTATGAGAGAG 1800
QY 1801 gatccaagcaatggccatggccgtaatcttgttaaatctagctgttlcagaatttccaag 1860
Db 1801 GATCCAAGCAATGCCATGCCGTAATCTTTGGTAAATCTAGCTGTTCAGAAATTTCAAAAG 1860
QY 1861 gaagcctatacagccgtagtatacataacaaggtctcctgatttccatgagaanaatcaag 1920
Db 1861 GAAGCCTATACAGCCGTAGTATATCATACAGGTCCTCGATTTTCATGAAGAAATCAAG 1920
QY 1921 gttaaagcttcctgctactttaactgaaccatcacttgccttttacttttcatcattgtt 1980
Db 1921 GTTAAAGCTTCTCTGCTACTTTAACTGACCATCATCATTGCTTTTACTTTTATATCATGTT 1980
QY 1981 agtgtcaacaaacaaacaaataactcctcttgaaacacacagttgatatatacatgatacca 2040
Db 1981 AGTTGTCAACAAACAAACAAATACTCTCTTGAAAAACACCAGTTGGATATACATGATGATCCA 2040
QY 2041 atgcttcaagaatggacggttgaaagacttggccagtttgcctggccagctcctcatatggaaaaa 2100
Db 2041 ATGCTTCAGAATGGACGGTGAAGACTGGCCAGTTTGGTGGCCAGTCTCATTTGAAAAA 2100
QY 2101 ccaccacaggttattctgtactgctctcctgtaggttccctctacactgycatgaatgggta 2160
Db 2101 CCACCACAGGCTTATTCTGTACTGCTCTCGTAGAGTTCCTTACCTGGCATGAATGGGTA 2160
QY 2161 gataatcaacaaaggtgtttttaatgttgaagttgttgcgttctcgtctatccatacacaa 2220
Db 2161 GATAATCAACAAGGTGTTTAAATGTTGAAGTTGTTGCTGTTCTATCCATACACAA 2220
QY 2221 gatccttatcttgacaaatlttctgctctgltcaatgctctgagatgaacacctgttccca 2280
Db 2221 GATCCTTATCTTGACAAATTTTGTGCTCTGTCATATGCTGTGATGAACACCTGTTCCCA 2280
QY 2281 gtccgaattgggagacatgacgaatcattgaaaaataaacttagaaaaatgaatgaagagcagt 2340
Db 2281 GTCCGAATTGGGAGACATGCGAATCATGGAATAAATACTTAGAAAAATGAATTGAAGAGCAGT 2340
QY 2341 attcagcacatgaattcatccacagctggaaccagtggtccgatttcttcattcttctgcta 2400
Db 2341 ATTTCAAGCACATGAATTCATCCACAGCTGGAACCAAGTGTCCGATTTCTTCATCTTCTGCTA 2400
QY 2401 gataaactgatacttttagttattagacctcctgtcatattgctgccaataagttaaccta 2460
Db 2401 GATAAACTGATACTTTAGTTATATAGACCTCCCTGTCATTTGCTGGCCAATAAGTTAACCTA 2460
QY 2461 ggtcaagcatcttttgaagccatggcattcaattataaatcgacttccacaaaaaacttggaa 2520
Db 2461 GGTCAAGCATCTTTTGAAGCCATGGCATCAATTTAATCGACTTCACAAAAAAGCTTGAA 2520
QY 2521 ggaatatcatgaccagcatggaagaacagccttcttgcatcatataattcatattgttttc 2580
Db 2521 GGAATCATGACCAGCATGGCAGAAACAGCCTTCTTGCATCATATATTCATTATGTTTTC 2580

QY 2581 cgcctaaccaaatacttaccctaattcatcatcaccaggtcctcgtggggttgggagatca 2640
Db 2581 CGCCTACCAAAATACTTACCCTTATTTTCATCATCACCAGGTCTCGGGGTTTGGAGATCA 2640
QY 2641 gtgcattatgccacaatggtcgtatagctcgtggtgagaccttgcgaagccttaattaaatcgt 2700
Db 2641 GTGCATTATGCCACAATGGCTAGATCTCGGGTGAGAGCCTTGCAAGCCTTAATTTAAATCGT 2700
QY 2701 tctcgaagccttagtaataagcaatccagatataatctggaactcccacgtcaccagatgat 2760
Db 2701 TCTCGAAGCCTTAGTAATAGCAATCCAGATATATCTGGGACTCCACGTCACACAGATGAT 2760
QY 2761 gaagttcgaatcaatcatcgggagtaagggtttagatcgctccaalttccgtggttaacact 2820
Db 2761 GAAGTTGATCAATCATCGGAGTAAAGGTTTAGATCGTCCAATTCCTGGGTTAAGACT 2820
QY 2821 ggtgtccaaaagctgcccatggggatccaacccagtcgaagtgcagaatcaacacag 2880
Db 2821 GGTGTCCAAAAGCTGCCCATGGGGATCCAAACCCAGTCCAAGTGAGAAATCAACACAG 2880
QY 2881 gctatgatcgaagtltgtaatcgtatgtcttcgcacacagagacgtcaagtttcttaaa 2940
Db 2881 GCTATGATCGAAGTGTGTAATCGTATGCTTCGCACACAGAGACGTCAAGTTCCTTACAA 2940
QY 2941 acattaacgagacgcttaccaaactaaaaagctlttcaacgagagctggtcttgcagtyg 3000
Db 2941 ACATTAACGGGAGCGCTTACCAACTAANAAGCTTTTTCACGAGAGAGCTGGCTTGCAGTGG 3000
QY 3001 gttgttgcagtygcagcgttccgggaatcagccttgcacaacagcctggtccttltttaa 3060
Db 3001 GTTGTTCAGTGGCAGCGTTCCGGGAATCATAGCTTTGCAACAAGCCTGCTCTTTTGTGAA 3060
QY 3061 ttaatgtaaaagacatggtgcacacattataactttaatgataaaacttgaagctccaaag 3120
Db 3061 TTAATGTTAAAGACATGGTGACACCATTTATACTTTAATGATTAACCTTGAGGCTCCAAG 3120
QY 3121 aaaagtcgttttccagaacglttcatatgatacattgcagctcttgcacagcagattgct 3180
Db 3121 AAAAGTCGTTTCCAGAAGCTTTCATGGAATGACATTGCAGCTCTTGTCACAGCATTTGCT 3180
QY 3181 agtataatagtttcaacgatttccagaagacacagaaatggttggagagactcaataacaagc 3240
Db 3181 AGTATATAGTTTTCAGCATTTTCAGAGAAGACACAGAAATGGTTGAGAGACTCAATACAAAGC 3240
QY 3241 ctlgcattcttccaaatgatcgtgtgtgttataagcagagagattgttlttagcctt 3300
Db 3241 CTTGCATTTCTTTCATATGATCTGTGTGCTGTATGGACAGAGAGATTGTTTAAAGCCTT 3300
QY 3301 ataaagtcctgtataaacaaggtgtcttcaaaagcttactcaattaccgaatcccaagtgt 3360
Db 3301 ATAAAGTCCTGTATPAAACAGGTGTCTTCAAAAGCTTTACTCATTTACCGAATCCAGTGT 3360
QY 3361 ctggtgtccttgagcgtggatttctacgaatcatctgcagtcatgagcaactatgttaca 3420
Db 3361 CTGGTGTCCTTGAGGCTGGATTTTCTACGAATCATCTGCAGTCATGAGCACTATGTTACA 3420
QY 3421 ttaacttaccctgcagcttacttactccacctgcactccatcactccttcttctct 3480
Db 3421 TTAACCTTACCCTGCAGCTTACTTACTCCACCTGCATCTCCATCACCTCTCTGTTCTTCT 3480
QY 3481 gcaacatctcagagttcttggatttctacgaatgtacaagacccaagaatgtcaaatatg 3540
Db 3481 GCAACATCTCAGAGTTCTTGATTTTCTACGAATGTACAAAGACCAAAAGATTGCAAAATATG 3540
QY 3541 ttgaaattatccgtgcctttccggccaacagcatatttggcagagacttgtgttaacagag 3600
Db 3541 TTTGAATTATCCGTGCTTTCCGCCAACAGCATTTATTTGGCAGGACTGTGTTAACAGAG 3600
QY 3601 ctggcgtcatatttagaccctgatgctgaagagactgttggattgcataaagaagtcatc 3660
Db 3601 CTGGCTGTCAATTTAGACCCCTGATGCTGAAGGACTGTTTGATTGCATTAAGAAAGTCATC 3660

QY 3661 aatatgtacacaattactctccagtcacgactcagaccgcggtactctgaccctcag 3720
|||||
Db 3661 AATATGTTACACAATTTACTCTCCAGTACGACTCAGACCCGCGGTACTCTGACCCCTCAG 3720
QY 3721 ataaagctcgaagtgcacatgtgtatctactctgattgtattatcatgynaactgta 3780
|||||
Db 3721 ATAAAGGCTCGAGTGGCCATGTTGTATCTACTCTGATTGTATATCATGGAACCTGTA 3780
QY 3781 cctcagctgtatgattttacagaacctcacaaatcaacgaggaagaccaatttgtatagcc 3840
|||||
Db 3781 CCTCAGCTGTATGATTTTACAGAACTCACAAATCAACGAGAGACCAATTTGTATAGCC 3840
QY 3841 actgatgattatgaagaagtgaagcggaagtatgataagccagaccglttgccatgccaatc 3900
|||||
Db 3841 ACTGATGATATGAAGTGAAGCGGAGATATGATAAGCCAGACCGTTGCCATGGCAATC 3900
QY 3901 gcagggacatcgttcctcctcaactaacaaggcctgycagtttcctcctcagctcaacgagt 3960
|||||
Db 3901 GCAGGGACATCGGTCCCTCAACTAACAGGCCCTGGCAGTTTCTCTCTCACGTCACGAGT 3960
QY 3961 ggcagggcaacaactaccttlttcagcagaaatcaagtcgaagccttttgaatctgttactt 4020
|||||
Db 3961 GGCAGGCAACACTACCTTTTCAGCAGAATCAAGTCGAAGCCTTTTGATCTGTCTACTT 4020
QY 4021 tgggtctcaaaaaatgcagatgnaaacgltctacagaagaagtgttacagatctctcagtc 4080
|||||
Db 4021 TGGTTCCTCAAAATGCAGATGAACAGTCTCTACAGAAGTGGTTTACAGATCTCTCAGTC 4080
QY 4081 ttgcagctaaaccggcctattagatctgctttatctctgtgtgtctgttcttgagtataa 4140
|||||
Db 4081 TTGCAGCTAAACCGGCTATTAGATCTGCTTATCTCTGTGTCTTGCTTTGAGTATAAA 4140
QY 4141 gggaaaaaagtgtttgaacgaatgaatagcttgaccctttaagaanaatcaaaagacatgaga 4200
|||||
Db 4141 GGGAAAAAGTGTTTGAACGAATGAAATAGCTTGACCTTTAAGAAATCAAAAGACATGAGA 4200
QY 4201 gcaaagcttgaagaagctatctctgggagcataggtgcccaggcaagaatggtacagcgga 4260
|||||
Db 4201 GCAAAGCTTGAAGAAGCTATCTTGGGAGCATAGGTGCCAGGCAAGAAATGGTACGGCGA 4260
QY 4261 agccgaggaagctcgagagaagcccatctggaagtgccttltggaagtcagaagaaatttg 4320
|||||
Db 4261 AGCCGAGGACAGCTCGAGAGAGCCCATCTGGAAGTGCCTTTGGAAGTCAAGAAATTTG 4320
QY 4321 agtgtgaggaagaatatactcaactggtgcgtcaaaaaacacagagaagcttgacaatatcaaga 4380
|||||
Db 4321 AGTGGAGGAAGAAGATATGACTCTACTGGCGCTCAAAAACACAGAGAAGCTTGACAAATCAAGA 4380
QY 4381 gcagagattgaaacaacgaagcactgattgtagaaccctggtctacagaagcaaacctaacc 4440
|||||
Db 4381 GCAGAGATTGAACACGAAAGCACTGATTGTGAAAACCTGGCTTACAGAAACCAACCTTAATC 4440
QY 4441 attttagatacatlagagattgtgttcagaccgttctgttaacgnaatccaagagagc 4500
|||||
Db 4441 ATTTTAGATACATTAGAGATTGTTGTTTCAGACCCGTTTCTGTAAACGGAATCCAAGAGAGC 4500
QY 4501 attcttggtgagtgctaaagtgctactacacagcatggtcgttaaccaaaagtgcagtt 4560
|||||
Db 4501 ATTCTTGGTGGAGTGTAAAGTGTCTACTACACAGCATGGCCCTGTAAACCAAGATGCAGTT 4560
QY 4561 tatctacaacactgttttgtctacacagagagccctgtttccaagtttccctgtaactctta 4620
|||||
Db 4561 TATCTTACACACTGTTTTCCTACACAGAGAGCCCTTGCTTCAAAAGTTTCCCTGAACTCTTA 4620
QY 4621 tttagaagaagagacagagcagtgctgtattatgcttcaggccttctccgacactgtlagc 4680
|||||
Db 4621 TTGGAAGAAGAGACAGAGAGCTGTGCTGATTATGCTTCAGGCTTCTCCGACACTGTAGC 4680
QY 4681 agtagcatcgtgtacaatacgtgtacaaccccaagtgtcccttacctactaatgagggcaa 4740
|||||
Db 4681 AGTAGCATCGGTACAAATACGCTACACCCAGTGTCCCTTTACCTACTAATGAGGCCAA 4740
QY 4741 aaatttgagattggaataacttltgcaggtttaaaatgcaggtaccaatgtcactatcc 4800
|||||

Db 4741 AACTTTGAGATTGGGAATTACTTTGCCAGGGTTAAATGACAGGTACCAATGTCACTATCC 4800
QY 4801 tcccttggtggcacatctcagaattttaatgaagaattcttaagacgttctctaagact 4860
|||||
Db 4801 TCCTTGGTGGGCACATCTCAGAATTTTATGAAGAAFTTCTTAAGACGTTCTTAAGACT 4860
QY 4861 atattgacatatgtctgaagaagaatctgnaattgagggaaacaacatltccctgatacagtc 4920
|||||
Db 4861 ATATTGACATATGCTGAAGAAGATCTGGAATTGAGGGAACAACATTTTCTGATCAGGTC 4920
QY 4921 caagatcgtgtttccaatctccatatgatatcttctgatactgtgaaaaatgaaggaaacac 4980
|||||
Db 4921 CAGGATCTGGTTTCAATCTCCATATGATTCTTCTGATACTGTGAAAAATGAAGGAACAC 4980
QY 4981 caaggaagatccctgaaatgttgattgatactaaigtacagaatlgccaaggttaccagacc 5040
|||||
Db 4981 CAGGAGGATCCTGAATGTGATTGATGATCTAATGTACAGAATTGGCAAGGTTTACCAGACC 5040
QY 5041 tctccagagcgattacacctggttgacagaacatgycaggcaagcactcagaaacgaagcaat 5100
|||||
Db 5041 TCTCCAGAGCGATTGACCTGGTTGCAGAAATGCGCAGGCAAGCACATGTCGAATATTGAGC 5100
QY 5101 catgctgaagctgcacagtgctcagtcacacacagcagcactgtgtgtgaatatlttagc 5160
|||||
Db 5101 CATGCTGAAGCTGCACAGTGTCTAGTCCACTCAGCACACACTTGTGCTGAATATTGAGC 5160
QY 5161 atgctgagagaccggaatatcttccctgtggtgagtgtgtaacatttccaagatatltcatct 5220
|||||
Db 5161 ATGCTGAGAGACCGGAATATATCTTCCGTGGGATGTGTAAATTTTCATCT 5220
QY 5221 aatgtttagaagaatctgcggtctcagatgtagtgtlatctccagatgaagaaggtatc 5280
|||||
Db 5221 AATGTTTAGAAGAACTCTGCGGTCTCAGATGATGTGTATCTCCAGATGAAGAAGTATC 5280
QY 5281 tgcctcgyaaaaatacttactgagtcagagacttgtggtatctcgyaaacaagcagctgct 5340
|||||
Db 5281 TGCTCTGGAATAATCTTACTGAGTCAGGACTTGTGGGATTACTGCAACAAGCAGCTGCT 5340
QY 5341 tccctctctatgctgctgcatglatgaagcagtttaatgaagtttcaaaagtlactatctc 5400
|||||
Db 5341 TCCTTCTATGGCTGGCATGTATGAAGCAGTTAATGAAGTTTACAAAGTACTTATTCCT 5400
QY 5401 attcatgaagctaaatcgygtagcgaagaacataccaacatltcagylaaactlcaagaa 5460
|||||
Db 5401 ATTCAATGAAGCTTAATCGGATGCAAGAACTATCCACAATTCATGTGTAACCTCAAGAA 5460
QY 5461 gcaatcagcaaaaatgtltcatcagagtaactggtggyagcgagtttggcacctatctc 5520
|||||
Db 5461 GCATTTCAGCAAAATTTGTTCAATCAGAGTACTGGCTGGAGCGGATGTTGGCACCTATTTT 5520
QY 5521 cgtgttggttttatatggaaccaagtccggygatttgatgnaaagaattgtttacaag 5580
|||||
Db 5521 CGTGTGTTTATGGAACCAAGTTCGGGATTTGGATGAACAAGAAATTTGTTACAG 5580
QY 5581 gagcctgcaataaccaacttgcagagatatctcacagattggaaggattttacgagaa 5640
|||||
Db 5581 GAGCCTGCAATAACCAAACTTGCAGAGATATCTCACAGATTGAGGGAATTTTACGGAGAA 5640
QY 5641 agatttggagagagatgtgtgtgaagtaatcaaaagactcctaaccctgtlagacaagtataa 5700
|||||
Db 5641 AGATTGGAGAGGATGTGTTGAAGTAATCAAGACTCTAATCTGTAGACAAGTGTAAA 5700
QY 5701 ttagatccctaacaaggcataatctcagaattacctatgtgagccatacttgcacacatat 5760
|||||
Db 5701 TTAGATCTTAACAAAGCATATATTCAGATTACCTATGTGAGCCATACTTTGACACATAT 5760
QY 5761 gagatgaaggaacagaatcacctaatctgcacaaaaatlacaatctcgtgcattcatgtac 5820
|||||
Db 5761 GAGATGAAGGACAGAATCACCTATTTTCGACAAAAAATTAACAATCTTGTGATTCATGTAC 5820
QY 5821 tgtacacccttactttagatggccgtgcccattggyggaacttcatgaacaattccaagaag 5880
|||||

Db	5821	TGTACACCCCTTTACTTTAGATGGCCGTGCCCATGGGGAACCTTCATGAACAATTCAAAGG	5880
QY	5881	aagaccattctgactacgctcctatgccttccctatatataaacaagggtcaatgtcact	5940
Db	5881		5940
Db	5881	AAGACCATTCTGACTACGCTTCATGCCCTTCCCTTAATTAATAACAAGGTCATGTCAC	5940
QY	5941	cataaagaagagatcatcttaacaccaatgtgaagtgcctattgagagacatgcagaaaaag	6000
Db	5941		6000
Db	5941	CATAAGAAGAGATCATCTTAACACCAATTTGAAGTTGCTATTGAGACATGCGAAAAAG	6000
QY	6001	acacagaggttgcatcttgcaacacatcagatccgcagaccccaaatgtctcagatg	6060
Db	6001		6060
Db	6001	ACACAGAGGTTGGCATTTGCAACACATCAGGATCCCGACAGACCCAAATGCTTCAGATG	6060
QY	6061	gtaccacaggatctgtagacacacacagtgaaatcaggggccttggaaagtgcacaggtt	6120
Db	6061		6120
Db	6061	GTACTCCAGGATCTGTAGGACCAACAGTGAATCAGGGGCTTTGGAAGTGCACAGGTT	6120
QY	6121	ttctgtctgaaatcacctagtgaaccccaagctcttcagacatcataataactgcgactc	6180
Db	6121		6180
Db	6121	TTCTGTCTGAATACCTAGTGAACCCAAAGCTCTTCAGACATCATATAAATGCGACCTC	6180
QY	6181	tgctttaagaagatttactaaaaagtgltgaagatgccttaagaaaaataagagcttaatt	6240
Db	6181		6240
Db	6181	TGCTTTAAAGATTTTACTAAAGGTGTGAAGATGCCCTTAAGAAAAAATAAGAGCTTTAATT	6240
QY	6241	ggcgcggttcaaaagagatcacaagaaggaattggggaactatctcgcccttaagaagc	6300
Db	6241		6300
Db	6241	GGCGCGGTCAAAAAGAGATATCAAGGGAATTTGGGAACCTATCTTCGCCCTTAAGAGCC	6300
QY	6301	ctacagccctagatcacagaaagtcctcagttatccaagccagtatgtgtccct	6360
Db	6301		6360
Db	6301	CCACAGCCCTAGATCACAGAAAGTCCCTCAGTTATCCAGCCAGTATGCTGTCCCT	6360
QY	6361	gccacagagattccttcagtcgaatgagcttgcgaatgtatctctaactgaatgca	6420
Db	6361		6420
Db	6361	GCCACAGAGATTCCTTCAGTCGATGAGCTTTCGCAAAATGGAATCTCTAATGATGCA	6420
QY	6421	ctgttttattcatctgcgaagaagccatgtattccaacatcgagtgltgaaaagatctattg	6480
Db	6421		6480
Db	6421	CTGTTTTATTCATCTGCAAAAGACCCATGTATTCAACATCGAGTGTGAAAAGATCTATTG	6480
QY	6481	gaaaccaacatggaaatggaattctgaaattatattcatctgaagaatgcagtgcccaag	6540
Db	6481		6540
Db	6481	GAAACCAACATGGAATGGAATTCGAAATTTATTATTCATTGAAGAATGCAGTGCCCAAG	6540
QY	6541	aaaatacaaatgtagattgttaacgcttgagaatcatgctatggttctaagtctcg	6600
Db	6541		6600
Db	6541	AAAAATCAAAATGATGATTGTAAACGCTTGAGAATCATGGCTTCTAATGTTCCG	6600
QY	6601	gtaacaagctgtatctttaagacatttaatgactcaagaaggtacatatatacttaacc	6660
Db	6601		6660
Db	6601	GTAACAAGCTGTATCTTTAAGACATTTTAATGACTCGAAAGGTACACTATACATTTACC	6660
QY	6661	attattatataccatagctaaaggttaaaaaattatcttaactttaagttcgatttttaatt	6720
Db	6661		6720
Db	6661	ATTATTTATACCATAGCTAAGGTTAAAAATTTATTCACCTTTAAGTTCGTATTTTAATT	6720
QY	6721	tataacacatttatagatctattggaccattttaaatgtagtaatgctattttaa	6780
Db	6721		6780
Db	6721	TATATCACCATTTATAGATTCAATTTTGGAACCCATTTTAAATGTAATGCTTATTTTAA	6780
QY	6781	aggtactaaaaaatatgtgaatgtttacctcgtgcgcgcagggccctc	6828
Db	6781		6828
Db	6781	AGGTACTAAAAAATATGTGAATGTTTACCTCGTGCGCGCAGGGCCCTC	6828
RESULT	2		
LOCUS	AB051558	4311 bp	mrna
DEFINITION	Homo sapiens mRNA for KIAA1771 protein, partial cds.		PRI 07-FEB-2001
ACCESSION	AB051558		
VERSION	AB051558.1	GI:12698086	

KEYWORDS	Homo sapiens cDNA to mRNA, clone:pj02581.
SOURCE	Homo sapiens
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (sites)
AUTHORS	Nagase,T., Kikuno,R., Hattori,A., Kondo,Y., Okumura,K. and Ohara,O.
TITLE	Prediction of the coding sequences of unidentified human genes. XIX. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro
JOURNAL	DNA Res. 7 (6), 347-355 (2000)
MEDLINE	21082932
REFERENCE	2 (bases 1 to 4311)
AUTHORS	Ohara,O., Nagase,T. and Kikuno,R.
TITLE	Direct Submission
JOURNAL	Submitted (22-NOV-2000) Osamu Ohara, Kazusa DNA Rsearch Institute, Department of Human Gene Research; 1532-3, Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:cdnainfo@kazusa.or.jp, URL:http://www.kazusa.or.jp/huge, Tel:81-438-52-3913, Fax:81-438-52-3914)
FEATURES	Location/Qualifiers
source	1..4311
gene	/organism="Homo sapiens"
CDS	/db_xref="taxon:9606"
	/clone="pj02581"
	/note="vector:phuescriptII SK plus"
	1..3910
	/gene="KIAA1771"
	<1..3910
	/gene="KIAA1771"
	/note="Start codon is not identified. pj02581 cDNA clone for KIAA1771 has a 1-bp deletion at the position 498 and a 110-bp insertion after the position 3119 of the sequence of KIAA1771, respectively."
	/codon_start=2
	/product="KIAA1771 protein"
	/protein_id="BAB21862.1"
	/db_xref="GI:12698087"
	/translation="KNEGNHDQGRNSLIHYVERLPNTYPNSSSPGGLGGS VHYATMARSAVRPASLNRSLSNSNPDISPTSPDDEVRISIGSKIDRSNWSV NTGPKAAPWGSNPBSAESTQAMDRSCNRMSSHTETSSFLQTLTGRLP TKLFHEEL ALQWVCSGSVRESALQOAMFEFFELMWKSMVHLYFNDKLEAPRKSRRPEREMDIAA LVSTIASDIVSRFQKTEMYERLNTSLAFNLDDLVMDRGFVFSLIKSCYQVSSKL YSLPNPSVLVSLRDLFLRICSHENHYVTLPCLSLTPASPSPSVSATSQSGFST NVQDQKIANMFEISVPRQHYLAGLVLTSLAVILDPDAEGLFGLHRYINMVHNL S SHSDPRYSDPQIKARVAMLYPLIGITIMEYVQLYDFTETHNQGRPICATDYES ESGMSISQTVAMAIAGTSVQLTRPGSFLTSRQHTTFSAESRSLICLMLVK NADETVLQKFTDLVQLNRLDLILYLCVSCFEYKGVKVFERNMSTLTKSKDMRAK LEEAILGISGARQEMVRRSGQLERSPSGSAFSCQENLRWRKDMTHMNRTEKLDKSR AEIEHEALIDGNLATEANLILDTLEIVQIVSVTESKESILGVLKYLHSMACNQ S AYLQHCFAQORALVSKPELLEFEETEQCADLCRLRLRHCS SIGTIRSHASASLYL LMRONFEIGNNFARVKMOTVMSLSSLVGTSQNFNEELRLRSKLTILTYAEEDLELRET TRPDQVODLVFNLMILSDTVKMKHEQEDPEMLIDLMYRIAKGYQTSPLRLTWLQNM AGHRSERNHAEAAQCLVHSAALVAEYLSMLEDRKYLVPVCVTFQNISNVLEESAVS DDVYSPDEEGICSGKFTESGLVGLLEQAAASFSMAGMYEAVNEVYKVLPIHEANRD AKRLSTIHGKLQEAFSKIVHQSSTGEMRFGTYFRVGFYGTGFGDLDEQEFVYKEPAT KLAEISRLEGFYGERGEDVVEVIKDSNPVCKKLDPNKAYIQITYVEPYFDTYEMK KEETILTPIEVAIEDMQTKQELAFATHODPADPKMLQMVLOGSGVTYVNOGPLVAVQ VFLSEIPSDPKLFRHNRKLRKCEKDEFTKRCEDALRKNSLIGPDQKEYQRELRNRYHR LKEALQPLINRKIPOLYKAVLPVTCHRDSFSRMSLRKMDL"

QY	2508	caaaacttggaaggaatcatgaccagcatgycagaacagcctcttgcatcatatat	2567
Db	1	CAAAAACCTTGAAGGAATCATGACCAGCATGGCAGAAACAGCCTTCTTGATCATATAT	60
Query Match	61.9%;	Score 4224;	DB 9; Length 4311;
Best Local Similarity	99.3%;	Pred. No. 0;	
Matches	4276;	Conservative	0; Mismatches 25; Indels 7; Gaps 3;

QY 2568 tcattatgtttccgcctaccaaatcttaacctaatlcatcaccaagtcctgggg 2627
|||||
Db 61 TCATTATGTTTCCGCCCTACCAAACTACTTACCCTTAATTCATCATCACACAGGTCCTGGGG 120
QY 2628 ttvggagagatcagtgcatatgccaacatggtcagatctgcgtgagacctgcaagcct 2687
|||||
Db 121 TTTGGGAGGATCAGTGCATTATGCCACAATGGCTAGATCTGCGGTGAGACCTGCAAGCCT 180
QY 2688 taattaaatcgltctcgaaacctagtaataagcaatccagatatatacttgggactccac 2747
|||||
Db 181 TAAATTAATCGTTCGGAACCTTAGTAATAGCAATCCAGATATATCTGGGACTCCAC 240
QY 2748 gtcaaccagatgataaagtlcgalcaatcatcgggagtaagggttagatcgctccaatc 2807
|||||
Db 241 GTCACCAGATGATGAAGTTCGATCAATCATCGGGAGTAAGGGTTAGATCGCTCCAATTC 300
QY 2808 ctgggttaaacactggtgtgccaaaaagctgccccatggggatccaaccccaagtccaagtgc 2867
|||||
Db 301 CTGGGTTAACACTGGTGTCCAAAAGCTGCCCATGGGGATCCAACCCCAAGTGC 360
QY 2868 agaatcaacacaggtatgatacgaagtgttaacgtatgtcttcgcacacagagacgtc 2927
|||||
Db 361 AGAATCAACACAGGCTATGGATCGAAGTGTAAATCGTATGTCTCGCACACAGACGCTC 420
QY 2928 aagtttctacaacatlaacgagcgcttacccaactaaaaagcttttcaacgaggagct 2987
|||||
Db 421 AAGTTCTTACAACAATTAACGGGACGCTTACCAACTAAAAAGCTTTTTCACGAGAGACT 480
QY 2988 ggcttgcaagtgggtgtgttcgagtcgagcgcttcgggaatcagctttgcaacaagcctg 3047
|||||
Db 481 GGCTTTGCAGTGGGTGTTTGACAGTGGCAGCGTTCGGGAATCAGCTTTGCCAACAGCCTG 540
QY 3048 gtctcttttgaatlaatggttaaagagcatggtgcacccattatacttcaatgataaact 3107
|||||
Db 541 GTTCTTTTGAATTAATGTTAAAGAGCATGGTGCACCATTTTATCTTAATGATAAAT 600
QY 3108 tgaagctccaagaaaaagtcgtttccagaacglttcatgtagacatgcaagctctgt 3167
|||||
Db 601 TGAAGCTCCAAGGAAAAGTCGTTTCCAGAAGCTTTTATGATGACATTCAGCTCTTGT 660
QY 3168 cagcaagatgctagtgatatagtttcaagatltcagaagacacagaaaatgltgagag 3227
|||||
Db 661 CAGCAGATTGCTAGTATAGTTTTCACGATTTCAAGAGACACAGAAATGTTGAGAG 720
QY 3228 actcaatacaagccttgcattcttctcaatgatctgtgtgttaatgagcagagatt 3287
|||||
Db 721 ACTCAATACAAAGCCTTGCATCTTCTCAATGATCTGTGTCTGTATGAGACAGAGATT 780
QY 3288 tgttttagcctataaagtctcgtctataaacaagtgltcttcaaaagctttactcattacc 3347
|||||
Db 781 TGTTTTAGCCTTAATAAGTCCCTGCTATAAACAGGTGCTCTTCAAAAGCTTTACTCATTTACC 840
QY 3348 gaatcccaagtgttctgtgtccttgagcgtggaatttctacgaatcatctgcagtcataga 3407
|||||
Db 841 GAATCCCAAGTGTCTGTGTCCTTGAGGCTGGATTCTTACGAATCATCTGCAGTCATGA 900
QY 3408 gcactatgtlaacatlaaaccttaacctgcagcttaacttaacctgcacatctccatcac 3467
|||||
Db 901 GCACTATGTTACATTAACCTTACCCTGCAGCTTACTTACTCCACCCTGCATCTCCATCAC 960
QY 3468 ttctgttctctcgacaacatctcagaggtctcggatttctacgaatgtacaagaccaaaa 3527
|||||
Db 961 TTTCTGTTCTTCTGCAACATCTCAGAGTTCGGATTCTTACGAATGTACAAGACCAAAA 1020
QY 3528 gattgcaaatatgtttgaattatccgtgtcctttccgccaacagcatatttggcagagact 3587
|||||
Db 1021 GATTGCAAAATATGTTGAATTAATCCGTGCTTTCCGCCAACACAGCATTAATTGGCAGGACT 1080
QY 3588 tgtgttaacagagctggcgtgtcatctttagaccctgatgctgaagagactgttggatgca 3647
|||||
Db 1081 TGTGTTAACAGAGCTGGCTGTCAATTTAGACCCCTGATGCTGAAGGACTGTTGGATTGCA 1140
QY 3648 taagaagatcatcaatatggtacacaatltactctccagtcacgactcagaccgcgta 3707

|||||
Db 1141 TAAGAAAGTCATCATATATGGTACACAATTTACTCTCCAGTCACGACTCAGACCCGGGTA 1200
QY 3708 ctctgacctcagataaaggtcagagtgycatgttgtatctaaccttgatgtattat 3767
|||||
Db 1201 CTCTGACCTCAGATAAAGGCTCGAGTGGCCATGTTGTATCTTACCCTGATGTGATTAT 1260
QY 3768 catggaactgtlaacctcagctgtatgattttacagaanaactcacatcaacagagaaagacc 3827
|||||
Db 1261 CATGGAACCTGTACCTCAGCTGTATGATTTTACAGAACTCACAAATCAACGAGAGAAGCC 1320
QY 3828 aatttgatagccactgatgatataagtgagagcggaagtatgataagccaagacct 3887
|||||
Db 1321 AATTGTATAGCCACTGATGATTAATGAAGTGAGAGCGGAGATATGATAAGCCAGACCGT 1380
QY 3888 tgcacatgccaatcgcaagggacatcggtccctcaactaacaaggcctgaggttctcctc 3947
|||||
Db 1381 TGCCATGGCAATCGCAGGGACATCGGTCCCTCAACTTAACAAGGCCCTGGCAGTTCTCCT 1440
QY 3948 cagctcaacgagtgycagggcaacacactacctttcagcagaatcaagtcgaagccttt 4007
|||||
Db 1441 CACGTCAACGAGTGGCAGGCAACACACTTACTTTTCACAGAGAATCAAGTCGAAGCCTTT 1500
QY 4008 gatgtctacttltgggtlctcaaaaaatgcagatgaaacagtlctacagaagtgtttac 4067
|||||
Db 1501 GATCTGTCTACTTTGGGTCTCTCAAAAATGCAGATGAACAGTTCTACAGAAGTGTCTTAC 1560
QY 4068 agatctctcagctcttgcaagctaaaccggtattagatcgtcttatctgtgtgtcttg 4127
|||||
Db 1561 AGATCTCTCAGTCTTGACGCTAAACCGGCTATTAGATCTGCTTATCTCTGTGTGCTCTTG 1620
QY 4128 ctltgagataaagggaanaagtglttgaacgaatgaatagcttgaccttaagaatc 4187
|||||
Db 1621 CTTGAGTATAAAGGGAANAAGTGTGTAACGAATGAATAGCTTGACCTTTAACAATC 1680
QY 4188 aaaaagacatgagagcaaaagcttgaagaagctatcttgyggagcatagtgccaaggaaga 4247
|||||
Db 1681 AAAAGCATGAGAGCAAAAGCTTTGAAGAAGCTATTCTTGGGAGCATAGGTGCCAGCAAGA 1740
QY 4248 aatgtaacgycgaagccgaggaacagctcgagagaagcccatctggaagtgccttgyaag 4307
|||||
Db 1741 AATGTAAGGGCGGAAGCCGAGGACAGCTCGAGAGAAGCCCATCTGGAAGTGCCTTTGGAAG 1800
QY 4308 tcaagaaaaatttgaagtgaaggaaalatgactcactgycgtlcaaaaacacagagaagct 4367
|||||
Db 1801 TCAAGAAAATTTGAGGTGAGGAAAAGATATGACTCACTGGCGTCAAAAACACAGAGAAGCT 1860
QY 4368 tgacaaatcaagagcagagatgaaacaagcaactgatatgynaacctggtctacaga 4427
|||||
Db 1861 TGACAATAACAAGACAGAGATTGAACACGAAGCAGATGTGATGGAACCTGGCTACAGA 1920
QY 4428 agcaaacctaatacatltttagatacatllagagatgtgtgttcagaccgttctgttaacyga 4487
|||||
Db 1921 AGCAAACCTAATCATTTTATGATACATTAGAGATTGTGTTACAGACCCTTGTGTAACGGA 1980
QY 4488 atccaaagagagacatctcgtgtgagtgctaaagtgtactacacagcatgycctgtaa 4547
|||||
Db 1981 ATCCAAGAAGAGCATTCTTGCTGAGTGTCTAAAGTGTCTACTACACAGCATGGCTGTAA 2040
QY 4548 ccaaaagtgcagtltatctcaacaactglttltgctacacagagagccttgyttcaaaagt 4607
|||||
Db 2041 CCAAGTGCAGTTTATCTCAACAACCTGTTTGTCTACACAGAGCCCTGGTTCAAAAGTT 2100
QY 4608 tccatgaactcttatttgaagaagacagagcagtgctgtatattatgctcaggtctct 4667
|||||
Db 2101 TCCTGAACCTCTTATTTGAAGAAGAGACAGAGCAGTGTGATTTATGCCCTCAGGCTCT 2160
QY 4668 ccgacactgttagcagtagcatcggtacaataagcgttcaacaccccaagtgcctcccttaact 4727
|||||
Db 2161 CCGACACTGTAGCAGTAGCATCGGTACAAATATCGGTACACAGCCAGTGCCTTTACCT 2220
QY 4728 actaatgaggcaaaacttggagattggyaataacttgcagggtlaaagtgcaggtacc 4787
|||||

Db 2221 ACTAATGAGCGCAAAACCTTTGAGATTGGGAATAACTTTGCCAGGGTTAAATGCAAGGTAAC 2280
QY 4788 aatgcactatcctccttggtgggcacatctcagaattttaatgaagaattcttaagacg 4847
Db 2281 AATGTCACCTATCCTCCTTGCTGGGCACATCTCAGAAATTTTAATGAAGAATTCCTTAAGACG 2340
QY 4848 ttctctaagaactatatgtgacatatgtctgaagaagatctggaattgagggaaacaacatt 4907
Db 2341 TTCTCTAAGACCTATATGTGACATATGCTGAAGAAGATCTGGAATTGAGGGAACAACATT 2400
QY 4908 tccgtatcaggtccaggaatcgggttttccaatctccaatatgtattcttctgtaactgtgaa 4967
Db 2401 TCCTGATCAGGTCAGGATCTGGTTTTCATATCTCCATATGATTTCTTCTGATACTGTGAA 2460
QY 4968 aatgaaggaacaaccaagagatccctgaaatgttgatgatactaattgtacagaattgcca 5027
Db 2461 AATGAAGGAACACACAGAGGATCCTGAAATGTTGATGATCTAATGTACAGAAATTGCCAA 2520
QY 5028 gggttaccagacctctccaga---gcgaltgacctggtgcagaaacatgycaggaagca 5084
Db 2521 GGGTTACCAAGACCTCTCCAGATCTGCAGATTGCACCTGGTTGCAGAACATGGCAGCGCA 2580
QY 5085 ctcagaacgaagcaatcatgctgaagctgcacagtgctcagtcactcagcagcacttgt 5144
Db 2581 CTCAGAACGAAGCAATCATGCTGAAGCTGCACAGTGTAGTCCACTCAGCAGCACTTGT 2640
QY 5145 tgcgtgaatatgtgagcatgctggaaggaaccggaatatcttcctgtggaatgtgtaacat 5204
Db 2641 TGCTGAATATTGTGAGCATGCTGGAGGACCGGAATATCTTCCTGTGGATGTGTACATT 2700
QY 5205 tcagaatatctcataatggttttagaagaatctgcgcgtctcagatgatagtgtatctcc 5264
Db 2701 TCAGAATATTTCATCTAATGTTTGAAGAATCTGCCGTCAGATGATGTGATCTCC 2760
QY 5265 agatgaagaaggtatctgctcctgaaaaatacttactgagtcaggaactgttggatctact 5324
Db 2761 AGATGAAGAAGGTATCTGCTCTGGAATAACTTTACTGAGTCAGGACTTGTGGGATTA 2820
QY 5325 ggaacaaagcagctgctcctcctcctatgctgcatgtataagcagttaatgaagtta 5384
Db 2821 GGAACAAGCAGCTGCTCTCTCTATGCTGCGATGTATGAAGCAAGTTAATGAAGTTTA 2880
QY 5385 caaagtactatctcatalcatgaagctaalcggaatgcagaagaaactataccacaattca 5444
Db 2881 CAAAGTACTTATTCCTATTCAATGAAGCTTAATCGGATGCAAGAACTATCCACAATTCA 2940
QY 5445 tggtaaacctcaagaagcattcagcaaaatgttcatcagagtaactggtggagcggt 5504
Db 2941 TGGTAACCTCAAGAAGCATTCAGCAAAATTGTTCAATCAGAGTACTGGCTGGAGCGGAT 3000
QY 5505 gtttgacacctatttcgtgtgtgttttatatggaaccaagtccgggatttgatgaaca 5564
Db 3001 GTTTGGCACCTATTTCGTGTGTTTATGGAACCAAGTCCGGGATTTGGATGAACA 3060
QY 5565 agaattgtttaacaaggagcctgcataaccaaaacttgcagagatatctcacagattgga 5624
Db 3061 AGAATTTGTTTACAAGGAGCCTGCATTAACCAAACTTGCAGAGATATCTCACAGATTGGA 3120
QY 5625 gggatttacaagaaagaatttgagaagatgtgttgaagtaatcaaaagctctaacc 5684
Db 3121 GGGATTTTACGAGAAAGATTGGAGAGATGTTGAAGTAATCAAAAGACTCTAATCC 3180
QY 5685 tgtagacaagtgtaaattagatcctaacaaggcatatatcagattacctaagttagacc 5744
Db 3181 TGTAGACAAGTGAATTAGATCCTAACAGGCAATATATTCAAGATTACCTATGTGAGCC 3240
QY 5745 atactttgacacatatgagatgaagacagaatcacctatttcgacaaaaattacaatct 5804
Db 3241 ATACTTTGACACATATGAGATGAAGGACAGAAATCACCTATTTCGACAAAAATTACAATCT 3300
QY 5805 tcgtcgattcatgtactgtacaccccttactttagatgagcggtgcccataggggaacttca 5864
Db 3301 TCGTCGATTCATGTACTGTACACCCCTTACTTTAGATGGCCGTGCCCATGGGGAACCTTCA 3360

QY 5865 tgaacaattcoaaagaagaccattctgactacgtctcagccttccctatatataaac 5924
Db 3361 TGAACAATTCAAAAGGAGACCATTCTGACTACGTCCTCATGCCCTTCTTATATTAACAAC 3420
QY 5925 aaggttcaatgtcactcataaagaagatcatctttaacaccaattgaaatgtctattga 5984
Db 3421 AAGGTCATGTCACTCATTAAGAAGAGATCATCTTAACCAATTTGAAGTGGCTATTGA 3480
QY 5985 ggcacatgcagaaaaaagacacagagagttggcatttgcacaacatcaggaatccgcagacc 6044
Db 3481 GGACATGCAGAAAAAGACACAGAGAGTTGGCATTTGCAACACATCAGGATCCCGCAGACC 3540
QY 6045 caaatgcttcagatgylactccaaggaatcgtlaggcaccaagltgaatcaggggcctt 6104
Db 3541 CAAAATGCTTCAGATGTACTCCAGGAGCTGTAGGCACACAGTGAATCAGGGGCCCTT 3600
QY 6105 ggaagtgcaggttttctgtctgaaatacctagtgaccccaagctctcagacatca 6164
Db 3601 GGAAGTTGCCAGGTTTTTCTGTGTAATACTAGTGACCCAAAAGCTCTTCAGACATCA 3660
QY 6165 taataactggaactctgctttaaagatttactaaagaagtgtgaagatgccttaagaa 6224
Db 3661 TAATAACTGCGACTCTGCTTAAAGATTTTACTAAAGGTGTGAAGATGCCCTTAAGAAA 3720
QY 6225 aaataagagcttaatttggcggttccaagaagaglatccaagggaaattg--yguaacta 6282
Db 3721 AAATAAGAGCTTAATTGGGCCGATCAAAAGAGTATCAAGGGAAGCTGGAGAGAACTA 3780
QY 6283 tcttcgcctlaaagaaggccctacagccctagatca--cagaagtcctcagttatccaa 6340
Db 3781 TCATCGCCTTAAAGAGGCCCTTACAGCCACTGATCAACAGAAAGATCCCTCAGTTATACA 3840
QY 6341 gccagtatgtctgtccctgcacacagagatccctcagtcgaatgaagccttcgcaaat 6400
Db 3841 GGCAGTATGCTTCACCTGCACACAGAGATTCCTTCAGTCGAATGAGCCCTTCGCAAAAT 3900
QY 6401 ggaatcctaactgaatgcactgttttatacatctgcagaagacatglatcaacatc 6460
Db 3901 GGATCTCTAAACGTGAATGCACCTGTTTATTCATCTGCAAGAGCCATGTATCAACATC 3960
QY 6461 gagtgtgaaaagatcttatggaaccaacatggaatggaattctgaaaattatattcat 6520
Db 3961 GAGTGTGAAGAAGATCTATTGGAAAAACAATGGAATGGAATTCTGGAATTATTATCAT 4020
QY 6521 tgaagaatgcagtgcccaagaanaatatcaaatgtagattgtttaagcgttgagaatcatg 6580
Db 4021 TGAAGAATGCAGTGGCCAAAGAAATATCAATGTAGATTGTTAACGCTTGAGAATCATGG 4080
QY 6581 ctatggttctaatgttcggtgaacaagcgtttatctttaaagacattttaatgactcaa 6640
Db 4081 CTATGGTTCTAATGTTCTGTGTAACAAGCTGTATCTTTTAAGACATTTTAATGACTCAA 4140
QY 6641 aggtacactatacatttaaccattatltatacagctaaaggtlaaaaaattatcaact 6700
Db 4141 AGGTACACTATACATTTTACATTATTATTAACATAAGCTAAGGTTAAAAATTTATTCACCT 4200
QY 6701 taagtcgtatttttaattatatacaccattatagattcatatttgaaccaatttaaa 6760
Db 4201 TAAGTTCGTATTTTAAATTATATTACCATTATAGATTCAATTTTGGAAACATTTTAAA 4260
QY 6761 ttagtaatgcttattttaaaggtactaaaaaatatgtgaatgtttac 6808
Db 4261 TGTAGTAATGCTTATTTAAGGTACTATTAAATATGTGAATGTTTAC 4308

RESULT 3
AX172880 AX172880 4144 bp DNA linear PAT 03-JUL-2001
LOCUS
DEFINITION Sequence 7 from Patent WO0142297.
ACCESSION AX172880
VERSION AX172880.1 GI:14597915
KEYWORDS

SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
AUTHORS	1 (bases 1 to 4144)
TITLE	Lu, P., Garman, J. D. and Candia, A. F.
JOURNAL	Clasp-3 transmembrane protein
FEATURES	Patent: WO 0142297-A 7 14-JUN-2001; Arbor Vita Corporation (US)
source	Location/Qualifiers
	1..4144
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/note="preliminary CLASP-3 cDNA"
CDS	<2..3610
	/note="unnamed protein product"
	/codon_start=1
	/protein_id="CAC43727.1"
	/db_xref="GI:14597916"
	/translation="RVNRSRLSNSNPDISGTPSPDDEVRSIIGSKGLDRSNSWNTGPKAAPWGSNPSPSAESTIQAMDRSCNRMSSHTESSFLQTLTGLPTKLFHEELALOMVCSGSVRESALQQAWEFFELMVKSMVHLYENDKLEAPRKSREPEREMDIALALSTIASDIVSRFQKDEWVERLNTSLAFNLDSLVMDRGFVSLIKSCYKQYSSKLYSLNPSVLYSLRLDFLRICSHHYVTNLPCSLTPPASPSVSSATSSSGSTNVQDQKIAMFELVSPFROQLHAGLVLELAVILPDAGLEGLHKVIMVHNLSSHDSDPKSDPQIKARVAMLYPLIGIMETVPQLYDFTETHNQRGPCIATDYESESGMSISQTVAMAIAGTSVPOLTRPGSEFLTSTSGRQHTTFSAESRSLICLLMWLKNADETVLOKWFDTLSVLQNLRLDLLLYLCVSCFEYKGVKVERMNSLTFFKSKDMRAKLEIALGISIGARQEMVRSRGQLERSPSGSAFSGQENLRMRKDMTHWRQNTKDKRAEIEHEALIDGNLATEANLILDTLEIVQTVSVTESKESILGGVLKVLHSMACNQSAVYLOHCFATORALVSKFPELLFEETEQCADLCRLLRHCSSSIGTIRSHPSASLYLLMRONEIIGNNFARVKMQVPMSSLVGTSONFNEFLRSLKTIITVAAEEDLELRETTFPDOYODLVENLHMILSDTYKMKHEQEDPEMLIDMRYIAKGYQTSPLRLTLWQNMAGKHSESNHAEAAQCLVHSALVAEYLSMLEDRKYLEPGCVTFONISSNVEESAVIDVSPDEEGICSGKYFTESGLVGLLEQAAASFMSAGMYEAVNEVYKVLIPHEANRDAKKLSIHGKLQEAFSKIVHQSTGWERMEGTFRVGYGTKEGDLDEQEFYKKEPAITKLAESHRLEGFYGEREGEDVVEVIKDSNPVDCKLDPNKAYIOITYVEPYEDTYEMKDRITYEDKNYNLRRFMCTPEFTLDGRAHGLHEOFKRTILTSHAEPYIKTRVNVTHKEEILTPRIEVAIEDMOKTQELAFATHQDPADPKMLQMVLOGSVGTTVNOGPLEVAQVFLSEIPSDPKLFRHNNKRLCFCDFTRCEDALRNKSLIGPVQKEYQRELGLKLSSP"
BASE COUNT	1252 a 829 c 890 g 1173 t
ORIGIN	
Query Match	60.4%; Score 4123; DB 6; Length 4144;
Best Local Similarity	99.9%; Pred. No. 0;
Matches 4136; Conservative	0; Mismatches 0; Indels 3; Gaps 1;
QY 2693	taaatcgltctcgaagccttagtaataagcaatccagatatatctgggactccacgtaac 2752
Db 6	TAAATCGTTCGGAAGCCTTAGTAATAGCAATCCAGATATATCTGGGACTCCACGTCAC 65
QY 2753	cagatgatgaagtcgatcaatcatcgggagtaagggttagatcgctccaattcctggg 2812
Db 66	CAGATGATGAAGTTCGATCAATCATCGGAGTAAGGTTTAGATCCGCTCCAAATTCCTGGG 125
QY 2813	ltaacactggtgtccaaaagctgccccatggygatccaacccagtccaagtgcagaat 2872
Db 126	TTAACACATGGTGGTCCAAAGCTGCCCATGGGATCCAAACCCAGTCCAAAGTCAGAAAT 185
QY 2873	caacacaggtatgatatcgaagttgtaatcgtatgtcttcgcacacacagagacgtcaagtt 2932
Db 186	CAACACAGGCTATGATCGAAGTGTGAATCGTATGTCTTCGCACACACAGAGACGTCAGATT 245
QY 2933	tcttacaacaattaaacyggaagcttaaccaactaaaagcttttcaacgagagctgycctt 2992
Db 246	TCTTACAACAATTAACGGGACGCTTACCAACTAAAAAGCTTTTTCACGAGAGCTGGCTT 305
QY 2993	tgcagtggtgtgttttgcagtggcagcgttcggyaatcagctttgcaacaagcctggttct 3052
Db 306	TGCAGTGGGTGTGTTGCAGTGGCAGCGTTCGGGAATCAGCTTTGCCAACAGCCCTGGTCT 365
QY 3053	tttttgaaattaatgtgtaaagagcatggtgcacacatttatactttaatgataaacttgagg 3112

Db 366	TTTTGAATTAATGGTAAAGAGCATGGTGACACCATTTTACTTTAATGATTAACCTTGAGG 425
QY 3113	ctccaaggaanaagtcggtttccagaacggttcatgtagacatltgcagctctgtcagca 3172
Db 426	CTCCAAGGAAAAAGTCGTTTCCAGAACGTTTCATGATGACATTTGCAGCCTCTTGTCAGCA 485
QY 3173	cgattgtagtgatatagtttcaacgatttccagaagagcacacagaanaatggttgagagactca 3232
Db 486	CGATTGCTAGTGATATAGTTTCAACGATTTCCAGAAGGACACAGAATAATGTTGAGAGACTCA 545
QY 3233	atacaagccttgcatctcttctcaatgatctgttctgtttagagagaggttggttt 3292
Db 546	ATACAGCCTTGCAATTTCTTCTCAATGATCTGTGTCTGTTATGACAGAGATTGTTT 605
QY 3293	ttagccttataaagtcctgtcatalaaacaggtgtcttcaaaagcttactcataccgaatc 3352
Db 606	TTAGCCTTATAAGTCCTGCTATTAACAGAGTGCTTCAAGCTTTACTCATTAACCGAATC 665
QY 3353	ccagtgctcgtgtcctcttgaggctggaatttctacgaatcatctgcagtcataagcact 3412
Db 666	CCAGTGTCTGTTGTCCTTGAGGCTGGATTTTTCTACGAATCATCTGCAGTCATGAGCAGCT 725
QY 3413	atgttacatlaaacctaccctgcagcttacttaccacctgcatactcaacctctcg 3472
Db 726	ATGTTACATTAAACTTACCTGCAGCTTACTTACTCCACCTGCATCTCCATCACCTTCTG 785
QY 3473	ttctcttcgaacatctcgaagttctgatttcttacgaaatgttacaagaaccaaaagattg 3532
Db 786	TTTCTTCTGCACACATCTCAGAGTCTCGATTCTTCTACGAATGTACAGACCAAAAGATTG 845
QY 3533	caaatatgtttgaattatccgtgccttccgccaacagcattatttgcaggaactgtgt 3592
Db 846	CAAAATATGTTGAATTAATCCGTGCCCTTCCGCCAACAGCATTATTGCGCAGCACTTGCT 905
QY 3593	taacagaagctggtcattttagacccttgatctgaaggactgtttggaattgcataaga 3652
Db 906	TAAACAGAGCTGGCTGTCATTTTAGACCCTGATGCTGAAGAGACTGTTGGATTGCATAGA 965
QY 3653	aagtcataaatatgttacacaaatttactctccagtlcaagactcagaccgcggtactctg 3712
Db 966	AAGTCATCAATATAGGTACACAAATTTACTCTCCAGTCACGACTCAGACCCGCGTACTCTG 1025
QY 3713	accctcagataaagctcgaagtcgagtcgcatgttgtatctactctgatttgtlatatcatg 3772
Db 1026	ACCCTCAGATAAAGGCTCGAGTGCGCCAGTGTGTATCTTACCTCTGATGGTATATCATGG 1085
QY 3773	aaactgtacctcagctgtatgattttacagaaaactcaaatcaacgaggaagaccaattt 3832
Db 1086	AAACTGTACTCTCAGCTGTATGATTTTACAGAAACTCACAATCAACGAGGAAGACCAATTT 1145
QY 3833	gtatagccaactgatatatgaagaagtgaagcgygaagtatgataagccagaccgttgcca 3892
Db 1146	GTAATGCCACTGATGATTATGAAGTGAGAGCGGAAGTATGATAGCCAGACCGTTGCCA 1205
QY 3893	tygcaatcgcaagggacatcggtccctcaactaaacaagcctgycagtttccctcctcaagt 3952
Db 1206	TGGCAATCGCAGGACATCGGTCCCTCAACTAACAAAGGCTGGCAGTTTCTCTCACGCT 1265
QY 3953	caacgagtcgcaagcgaacacactacccttctcagcagaataccaagtcgaagcctttgatct 4012
Db 1266	CAACGAGTGGCAGGCAACACACTACCTTTTCAGCAGAAATCAAGTCGAAGCCTTTTGATCT 1325
QY 4013	gtctacttgggtctcaaaaatgcagatgaaaaacagttctacagaagtggtttacagatc 4072
Db 1326	GTCCTACTTTGGGTCTCAAAAATGCAGATGAACAGTCTTACAGAACTGGTTTACAGATC 1385
QY 4073	tctcagctcttcgagctaaaccggtatctagatctgcttatactctgtgtgtctgtcttg 4132
Db 1386	TCTCAGTCTTGCAGCTAAACCGGCTATTAGATCTGCTTATCTCTGTGTCTGCTTGTG 1445
QY 4133	agtataaagggaaaaaagtggtttgaacgaaatgtaatagcttgacctttaagaatacaaaag 4192

Db 1446 AGTATAAGGGAAAAAGTGTGACGAATGAATAGCTTGACCTTTAAGAAATCAAAAG 1505
QY 4193 acatgagagcaaaagcttgaagaagctatctcttggagcatagtgccaggaagaatgg 4252
Db 1506 ACATGAGAGCAAAAGCTTGAAGAAGCTATTCTTGGGAGCATAGGTGCCAGGCAAGAAATGG 1565
QY 4253 tacgycgaagccgaagcagctcagagaagcccatctggaagtgccttggaaagtcaag 4312
Db 1566 TACGGCGAAGCCGAGGAGCACTCGAGAGAAGCCCATCTGGAAGTGCCCTTGGAAAGTCAAG 1625
QY 4313 aaaattgagtgaggaagaatatgactcactgycgtcaaaaacacagagaagcttgaca 4372
Db 1626 AAAATTTGAGGTGAGGAAGATATGACTCAGTGGCGTCAAAACACAGAGAAGCTTGACA 1685
QY 4373 aatcaagagcagagattgaacacgaagcactgattgtaagaaacctggtacagaagcaa 4432
Db 1686 AATCAAGAGCAGAGATTGACACGAGCAGCTGATTGGAACCTGGCTACAGAAGCAA 1745
QY 4433 acctaatcattttagatcattagagattgltgttcaagaccgttctgtaacggaatcca 4492
Db 1746 ACCTAATCATTTTAGATACATTAGAGATTGTTGTTCAGACCGTTCGTGTACGGAATCCA 1805
QY 4493 aagagaagcattctgtgtgagtgctaaaagtgctactacacagcatgycctylaaccaaa 4552
Db 1806 AAGAGAGCATTTCTGGTGAGTGCTAAAAGTGCTACTACACAGCATGGCCTGTAACCAA 1865
QY 4553 gtcagtttatactacaacactgtttgtctacacagagacccttggttcaaaqtttcctg 4612
Db 1866 GTCGAGTTTATCTACAACACTGTTTGTGTACACAGAGACCCCTTGTTCAAGTTTCCTG 1925
QY 4613 aaccttatitgaagaagacacagagcagtggtgtgtatltatgctcaggtcttcocgac 4672
Db 1926 AACCTTATTTGAAGAAGACAGAGCAGTGTGCTGATTTATGCCCTCAGGCTTCCTCCGAC 1985
QY 4673 actgtagcagtagcatcggtacaataacggtcacaccccagtgccctcccttaactactaa 4732
Db 1986 ACTGTAGCAGTAGCATCGGTACAAATACGGTCAACCCCCAGTGCCCTTACCCTACTAA 2045
QY 4733 tgaagcaaaaactttgagattggygaataactttgccaggggttaaaatgcaggtaccaatgt 4792
Db 2046 TGAAGCAAAACTTTGAGATTGGGAATTAACCTTGGCCAGGGTTAAAATGCAGGTACCAATGT 2105
QY 4793 cactatccctccttgtyggcacatctcagaatlttaatgaagaatcttaagacgltctc 4852
Db 2106 CACTATCCTCCTTGGTGGCACATCTCAGAAATTTAATGAAGAATTCTTAAGACGTTCTC 2165
QY 4853 taaagactataitgacatatgctgaagaagatctgaaatitgagggaaacaacatttcctg 4912
Db 2166 TAAAGACTATATTGACATFATGCTGAAGAAGATCTGGAATTTGAGGAAACAATTTCCTG 2225
QY 4913 atcaggtccagatctgtgtttcaatctccatatgattcttctgatactgtgaaaaatga 4972
Db 2226 ATCAGGTCCAGGATCTGGTTTCAATCTCCATATGATTTCTTGATACTGTGAATAATGA 2285
QY 4973 aggaacaccaggygatccctgaaaaatgttgattgatactaaatgtlacagaatgtgccaaaggtt 5032
Db 2286 AGGAACACCAGGAGGATCCTGAAATGTTGATTGATCTAATGTACAGAAATTTGCCAAGGGTT 2345
QY 5033 accagaccctctccaga--gcgaattgacctgttcagaaacatgycaggaagcagcactcag 5089
Db 2346 ACCAGACCTCTCCAGATCTGGATTTGACCTGGTTGCAGAAACATGGCAGGCAAGCACTCAG 2405
QY 5090 aacgaagcaatcatgtctgaagctgcacagtgctctagtcgaactcagcagcactgttgctg 5149
Db 2406 AACGAAGCAATCATGTCTGAAGCTGCACAGTGTCTAGTCCACTCAGCAGCACACTGTGCTG 2465
QY 5150 aatatitgacatgctgagagaccggaatatatcttcctgtggyatgtgtaaacatttcaga 5209
Db 2466 AATATTTGAGCATGCTGAGAGACCGGAATATCTTCCTGTGGGATGTGTAACTTTTCAGA 2525
QY 5210 atattcatctaatgttttaagaagaatctgcggtctcagatgatgtgtatatctccagatg 5269
Db 2526 ATATTTCATCTAATGTTTTGAAGAATCTGCGGCTCTCAGATGATGTGATATCTCCAGATG 2585

QY 5270 aagaaglatctgctctggaataacttactgagtcaggaacttgyggaattacttggaaac 5329
Db 2586 AAGAAGTATCTGCTCTGGAANAATACTTACTGAGTCAAGGACTTGTGGATTTACTGGAAC 2645
QY 5330 aagcagctgtcttcctctctatggtcgtgcagtgtatgaagcagttaatgaagtttacaag 5389
Db 2646 AAGCAGTGTCTCCTCTCTATGGCTGGCATGTATGAAAGCAGTTAATGAAGTTTACAAAG 2705
QY 5390 tacttattccatttcatagaagctaactcggyatgcgaagaactataccacaattcatgta 5449
Db 2706 TACTTATTTCTATTTCATGAAGCTAATCGGGATGCAAGAATACTATCCACAATTTCATGTA 2765
QY 5450 aactcaagaagcaltcagcaaaaattgltcatcagagtaactgycgtggaagcggaatttg 5509
Db 2766 AACTTCAGAAGCATTCAGCAAAATGTTTCATCAGAGTACTGGCTGGAGCGGATGTTTG 2825
QY 5510 gcacctatttcgtgtgtgttttatgnaaccaagttcggggatttggatgaacaagaat 5569
Db 2826 GCACCTATTTTCGTGTGTGTTTATGGAACCAAGTTGCGGGAATTTGGATGAACAAGAAT 2885
QY 5570 ttgtttacaagagcctgcataaaccaaacttgcagagatatctcacagattggaaggtat 5629
Db 2886 TTGTTTACAAGGAGGCTGCAATTAACCAAACTTGCAGAGATATCTCACAGATTGGAGGAT 2945
QY 5630 tttaaggaagaaagatttggagagatgtgtgtgaagtaatcaaaagactctaalcctgtag 5689
Db 2946 TTTACGGAGAAAGATTGGAGAGATGTGTTGAAGTAAATCAAGACTCTAATCTGTAG 3005
QY 5690 acaagtgtaaatlagatccctaacaagcatatatlcagaattaactatgtgnaagcatact 5749
Db 3006 ACAAGTGTAAATTAGATCCTAACAGGCATATATTAGATTACCTATGTGAGCCATACT 3065
QY 5750 ttgacacatatgagatgaagagacgaatcacctatttcgacaaaaattacaatttcgtlc 5809
Db 3066 TTGACACATATGAGATGAAGGACAGAATCACCTATTTCGACAAAATTTACAATCTTCGTC 3125
QY 5810 gattcatgtaactgtacaccccttaacttagatgycggtgcccatggggaactcattgaac 5869
Db 3126 GATTCAATGTAAGTGTACACCCCTTACTTTAGATGGCCGTGCCCATGGGAACCTCATGAAC 3185
QY 5870 aattcaaaaaggaagaccattctgactacgtctcatgccccttccttatataatlaaacaaggy 5929
Db 3186 AATTCAAAAGGAAGACCATTCTGACTACGTCATGCGCTTCCCTTATATTAAACAAGGG 3245
QY 5930 tcaatgtcaactcataaagaagagatcatctlaacaccaaattgaaagttgctatgaagaca 5989
Db 3246 TCAATGTCACTCATAAAGAAAGATCATCTTAACACCAATTGAAGTTGCTATTGAGGACA 3305
QY 5990 tgcagaaaaagacacagagttgcaatttgcaacacatcaagatcccgagaccocaaaaa 6049
Db 3306 TGCAGAAAAAGACACAGAGTTGGCATTTGCAACACATCAAGGATCCCGCAGACCCCAAAA 3365
QY 6050 tgcttcagatgtaactccaggtatctgtaggcacacacagtgaaatcaggggccttggaaag 6109
Db 3366 TGCTTCAGATGGTACTCCAGGAGATCTGTAGGCACACAGTGAATCAGGGCCTTTGGAAG 3425
QY 6110 ttgcccaggttttctgtctgaaataactagtgaccccaagctcttcagacataataa 6169
Db 3426 TTGCCAGGTTTTCTGTCTGAATTAACCTAGTGACCCCAAGCTCTTCAGACATCATATA 3485
QY 6170 aactgcgaactgtcctttaaagatttacttaaaaggtgtgaagatgccttaagaaaaata 6229
Db 3486 AACTGCGACTCTGCTTTAAAGATTTTACTTAAAGGTGTGAAGATGCCCTTAAGAAAAATA 3545
QY 6230 agagcttaattggyccggttcaaaagagtaacaaaggaattggygaactatctcgc 6289
Db 3546 AGAGCTTAATTGGCCGGTTCAAAAGAGTATCAAAAGGGAATTGGGAAACTATCTTCGC 3605
QY 6290 ctlaaagggccctacagccctagatcacagaagaatccctcaqttalccaagccagtat 6349
Db 3606 CTTAAAGAGGCCCTACAGCCCTAGATCACAGAAAGTCCCTCAGTTATCCAAAGCCAGTATT 3665

QY	6350	gcttgccccctgcacagagatccttcagtcgaatgagcttcgcaaaatgcatctcta	6409
Db	3666	GCTTGCCCCCTGCACAGAGATTCCTTCAGTCGAATGAGCTTTCGCAAAATGGATCTCTTA	3725
QY	6410	aactgaatgcacttgtttatcatctgcgaagaagccatgatatcaacatcgagtgtgaa	6469
Db	3726	AACTGAATGCACCTGTTTATTTCATCTGCGAAAGAGCCATGTATTCAACATCGAGTGTGAA	3785
QY	6470	aagatctattgnaaaccaacatggaatctctgnaatttatattcatttggaagatg	6529
Db	3786	AAGATCTATTGGAACCAACATGGAATGGAATTCCTGGAATATTATTTCATGAGAATG	3845
QY	6530	cagtgcccaagaanaatatcaaatgtagatgtttaacgcttgagaatcatgctatggtt	6589
Db	3846	CAGTGCCCAAGAAATATCAAAATGTAGATTGTTAACGCTTGAGAATCATGGCTATGCTTT	3905
QY	6590	ctaagtctcggttaacaagctgttatcctttaaagacatttlaatgaactcaaaaggtacact	6649
Db	3906	CTAATGTCGGGTAAACAAGCTGTATTCTTTTAAGACATTTTAATGACTCAAAAGGTACACT	3965
QY	6650	atacattaccattattatataccatagctaaggtlaaaaaattatcaactttaaagttcgt	6709
Db	3966	ATACATTTACCATTATTATTATACCAATAGCTTAAGTTAAAAATTTATTCACTTTTAAGTTGGT	4025
QY	6710	atttttaattatatacaccattatagattcatcttggaccattttaaatgtagtaac	6769
Db	4026	ATTTTAAATTTATATCACCATTTATAGATTCAATTTTGACCCATTTTAAATGTAGTAAT	4085
QY	6770	gcttattttaaaggtactaaaaaataatgtgaatgtttaccctcgtgcgcgccagggcctc	6828
Db	4086	GCTTATTTTAAAGGTACTAAAAAATATGTGAATGTTTACCCTCGTGGCGCCAGGGCCCTC	4144

RESULT 4

AXI72932	AXI72932	4144 bp	DNA	linear	PAT 03-JUL-2001
LOCUS	AXI72932	Sequence 59 from Patent WO0142297.			
DEFINITION	AXI72932				
ACCESSION	AXI72932.1	GI:14597917			
VERSION					
KEYWORDS					
SOURCE		human.			
ORGANISM		Homo sapiens			

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 4144)
AUTHORS Lu, P., Garman, J.D. and Candia, A.F.
TITLE Clasp-3 transmembrane protein
JOURNAL Patent: WO 0142297-A 59 14-JUN-2001;
Arbor Vita Corporation (US)

FEATURES
source location/Qualifiers
1..4144
/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="human Clasp-3 cDNA"
197..3610
/note="human Clasp-3"
/codon_start=1
/protein_id="CAC43728.1"
/db_xref="GI:14597918"

CDS
SVRESALQOAWFEFELWKSVMVHLYENDKLEAPKSRPERFMDIAALVSTIASDI
VSREOKDTEMERLNLFLNDLISVMDRGFEVSLIKSCYKQVSSLSLPLNPSVL
VSLRDLFLRICSHHYVTNLPCSLTPPASPPSVSSATSSQSGFSTNVQDOKIAN
MELSVPRQOHYLAGLVITELAVILDPDAEGLFGLHKKVINMVHNLSSHSDPRYS
DPOIKARVAMLYPLIGIMETVPOLYDFETHNQRGPICATDDSESGMSISQT
VAMAIAGSTVPQLTRPGSFLITSTSGRQHTTSAESSRSLICLLWVILKNADETVLQK
WETDLSVLQNLRLDLILYCVSCEFYEGKGYEERMSLTFKKSKDMRAKLEAEILGSI
GARQEMVRRSRQLESPSGSAFGSOENLRKDMTHWRONTKLDKSRAEIEHEALI
DGNLATEANLILIDPLEIVQTVSVTESKESTIGVLKVLHSMACQSAVYLQHCFA
TORALVSKPELLEFEETEQCADLCRLRLRHCSSSIGTIRSHPSASLYLLMRONFEIG
NNFARVKQVPMSSLISLVGTSQNFNEFLRSLIKTILTYAEEDLELRETTFPDQVQDL
VENLHMLISDTVKMKEHQEDPEMLIDLMYRIAKGYQTSPLRLTFWLNMAKGHSERN
HAEAQCLVHSAALVAEYLSMLEDRKYLPGCVTFQNISSNVLEBSAVSDDVVSPDEE

QY	2693	taaatcgltctcgaagccttagtaataagcaatccagatatatctgggaactcccaagtcac	2752
Db	6	TAAATCGTTCGGAAGCCTTAGTAATAGCAATCCAGATATATCTGGGACTCCACGTCAC	65
QY	2753	cagatgataagtcgatcaatcatcgggagtaagggtttagatcgctccaattcctggg	2812
Db	66	CAGATGATGAAGTTGATTCATCATCGGAGTAAGGTTTAGATCGCTCCAATTCCCTGGG	125
QY	2813	ttaacactggtgtccaaaagctgccccatggggtoccaaccccgatccaagtgcagaat	2872
Db	126	TTAACACTGGTGGTCCAAAAGCTGCCCATGGGATCCAACCCAGTCCAAGTCAGAAAT	185
QY	2873	caacacaggtatgatacgaagttgtaatcgtatgtcttcgcacacagaagaagtcgaagt	2932
Db	186	CAACACAGGCTATGGATCGAAGTGTATATCGTATGTCTTCGCACACACAGACGTCAACTT	245
QY	2933	tcttaacaacattaaagggagcgttaaccaataaaaagccttltcaagaggagctggtc	2992
Db	246	TCTTACAACATTAAACGGGACGCTTACCACACTAAAAAGCTTTTTCACAGAGGAGCTGGCTT	305
QY	2993	tgcagtggtgtgttcgcaagtcggcagcgttcgggaaatacagctlttgcaacaagcctggtct	3052
Db	306	TGCAGTGGTGTGTTCAGAGTGCGACGCTTCGGGAATCAGCTTTGCAACAAGCCTGGTCT	365
QY	3053	ttttgtaattaatgtaaaagagcatggtgcacacattatactttaatgataaactggag	3112
Db	366	TTTTTGAATTAAATGTTAAAGAGCATGGTGCACCAATTATCTTTAAATGATAAACTTGAGG	425
QY	3113	ctccaagaaaagtcggttttccagaacggttcatagtatgacattgcagctctgtcagca	3172
Db	426	CTCCAAGGAAAAGTCGTTTTCAGAACGTTTCATGGATGACATTGCAAGCTCTGTCAAGCA	485
QY	3173	cgattgtagtataagtttcaacgatttcaagaagagacacagagaatggttgagagactca	3232
Db	486	CGATTGCTAGTATAGATTTCACGATTTCAGAAAGACACAGAAATGGTTGAGAGACTCA	545
QY	3233	atacaagccttgcatcttctcgaatgatacgtgtgtctgttatgagacagaggattgltt	3292
Db	546	ATACAAGCCTTGCAATCTTCTCAATGATCTGTGTCTGTATGACACAGGATTTGTTT	605
QY	3293	ttagcctataaagctcgtctataaacaaggtgtcttcaaaagcttactcattaccgaatc	3352
Db	606	TTAGCCTTATAAGTCTCTGCTATAAACAGGTCCTTCAAAAGCTTTACTTCATTACCGAATC	665
QY	3353	ccagtgctcgtgtccttgagcgtgatttcttaagaaatcatctgcagtcatgagcact	3412
Db	666	CCAGTGTCTGTCCTTGAGGCTGGATTTCTACGAATCATGTCAGTGCATGAGCAGCT	725
QY	3413	atgtacattaaacttaccctgcagcttacttactccacctgcatactccatccactctg	3472
Db	726	ATGTTACATTAACTTAACCTGCAGCTTACTTACTCCACCTGCATCTCCATCACCCTTCG	785
QY	3473	tttctctgcaaatctcagagttctgatatcttctacgaatgtataagaaccaaaagatg	3532
Db	786	TTTCTTCTGCACAATCTCAGAGTTCTGGAATTTTCTACGAATGTACAGAACCACAAAGATTG	845
QY	3533	caaatatgtttcaattatccgtgccttccgcgaacagcatlataattggcaggactgtgt	3592
Db	846	CAATATATGTTGAATTATCCGCTTCCTCCGCCAACAGCATTAATTGGCAGGACTGTGT	905

QY 3593 taacagagctggctgtcatcttagaccctgatgctgaagactgtttggtatgcataaga 3652
|||||
Db 906 TAACAGAGCTGGCTGTCACTTTAGACCCTGATGCTGAAGACTGTTGGATTGCATAAGA 965
QY 3653 aagtcataatatgtgtacacaatttactctccagtcacgactcagaccggcggtactctg 3712
|||||
Db 966 AAGTCATCATATGTGTACACAATTTACTCTCCAGTCACGACTCAGACCCGGGTACTCTG 1025
QY 3713 accctcagataaaggctcgagtgccatgttgtatctactctgtattgttataatg 3772
|||||
Db 1026 ACCCTCAGATAAAGGCTCGAGTGGCCATGTTGATCTACCTCTGATTGGTATTATCATGG 1085
QY 3773 aaactgtacctcagctgtatgatttacagaactcacaatcaacgagaagcaaatlc 3832
|||||
Db 1086 AAACGTACTCTCAGCTGTATGATTTTACAGAAACTCACAAATCAACGAGGAAGACCAATT 1145
QY 3833 gtatagccactgatgatattagaaagttagagcggaagtatgataagccagaccgttgca 3892
|||||
Db 1146 GTATAGCCACTGATGATTATGAAGTGAGAGCGGAAGTATGATAAGCCAGACCGTTGCCA 1205
QY 3893 tggcaatcgcaaggacatcgtgccctcaactaacaaggccctggcagtttcctccacgt 3952
|||||
Db 1206 TGGCAATCGCAGGAGCATCGGTCCCTCAACTAACAGAGCCCTGGCAGTTCCCTCCACGT 1265
QY 3953 caacgagtggcaggcaacacacttactcttcagcagaatcaagtgcgaagcctttgatct 4012
|||||
Db 1266 CAACGAGTGGCAGGCAACACACTTACTTTTCAGCAGAAATCAAGTCGAAGCCTTTTGATCT 1325
QY 4013 gtctactttgggtctcaaaaatgacagatgaaacagttctacagaagtgtttacagatc 4072
|||||
Db 1326 GTCTACTTTGGGTTCTCAAAAATGCAGATGAACAGTTCTACAGAAAGTGTTTACAGATC 1385
QY 4073 tctcagttcttgcaagtaaacccggtactatagatctgcttatactctgtgtcttgcttg 4132
|||||
Db 1386 TCTCAGTCTTGCACTAAACCGGCTATTAGATCTGCTTATCTCTGTGTCTTGCTTG 1445
QY 4133 agtataaaggaaaaaagtgtttgaacgaatgaatagcttgacccttaagaatcaaaag 4192
|||||
Db 1446 AGTATAAAGGAAAAAGTGTGTAAGCAATGAATAGCTTGACCTTTAAGAATCAAAAG 1505
QY 4193 acatgtagcgaagcttgagaagcttattcttgaggacatagtgccagcgcaagaatg 4252
|||||
Db 1506 ACATGAGAGCAAAAGCTTGAAGAAGCTATTCTTGGGAGCATAGGTGCCAGGCAAGAAATGG 1565
QY 4253 tacggcgaaagccgaggagacagctcgagagaagcccatctggaagtgcctttggaagtcaag 4312
|||||
Db 1566 TACGGCGAAGCCGAGGACAGCTCGAGAGAAAGCCCATCTGGAAGTGCCCTTTGGAAGTCAAG 1625
QY 4313 aaaattgagtgaggaaagatatgactcaactgycgtcaaaacacagagaagcttgaca 4372
|||||
Db 1626 AAAATTGGAGGTGAGGAAAGATATGACTCACTGGCGTCAAAACACAGAGAAGCTTGACA 1685
QY 4373 aatcaagagcagagatlgaaacagaaagcactgattgatygaaacctggctacagaagcaa 4432
|||||
Db 1686 AATCAAGAGCAGAGATTGAACACGAAAGCACTGATTGATGAAACCTGGCTACAGAAGCAA 1745
QY 4433 accataatcatcttagatacatatagagattgttttcagaccgtttctgttaacgnaatcca 4492
|||||
Db 1746 ACCTAATCATTTTAGATACATTAGAGATTGTTGTTTCAGACCGTTCGTGAACGGAATCCA 1805
QY 4493 aagagagcatctcttggtgtgagtgctaaaaagtgtctactacacagcatggcctgtaaccaaa 4552
|||||
Db 1806 AAGAGAGCATTTCTTGGTGGAGTGCTAAAGTGCTACTACACAGCATGGCCTGTAAACCAA 1865
QY 4553 gtgcagttatactacaacactgttttgtacacagagagacctgtgttcaaaagttcctg 4612
|||||
Db 1866 GTGCAGTTTATCTACAACACTGTTTGTCTACACAGAGAGCCCTTGTTTCAAAAGTTTCCCTG 1925
QY 4613 aactcttattgaagaagagacagagcagtgctgtattatgctcaggtctctccgac 4672
|||||
Db 1926 AACTCTTATTGGAAGAGACAGAGCAGTGTGCTGATTATGCCCTCAGGCTTCTCCGAC 1985

QY 4673 actgtagcagtagcatcgtgtacaatacgytcaacacccacgtgcctcccttactactaa 4732
|||||
Db 1986 ACTGTAGCAGTAGCATCGGTACAATACGGTCAACACCCAGTGCCTCCCTTACCTACTAA 2045
QY 4733 tgaggcaaaccttgagattgygaataactttgcagggttaaaaaatgcaggtlaccaatgt 4792
|||||
Db 2046 TGAGGCAAAACTTTGAGATTGGGAATTAACCTTGGCAGGGGTAAATGCAGGTACCAATGT 2105
QY 4793 cactatcctccttggtlgygcacatctcgaatlttaatgaagaattcttaagacgttctc 4852
|||||
Db 2106 CACTATCCCTCCTTGGTGGGCACATCTCGAAATTTTAATGAAGAAATCTTAAGACGTTCTC 2165
QY 4853 taaagactatatgacatatgtctgaagaagatctgnaatlgagggaaacacattcctg 4912
|||||
Db 2166 TAAAGACTATATGACATATGCTGAAGAGATCTGGAATTGAGGGAACAACATTTCCCTG 2225
QY 4913 atcagttccagatctgtgttttcaatctccatatgattcttctgtatactgtgaaaatga 4972
|||||
Db 2226 ATCAGGTCCAGATCTGGTTTTCATATCTCCATATGATTCTTTCTGATACTGTGAATAATGA 2285
QY 4973 aggaacacccaggagatccctgnaatgttgattgatctaatgtlacagaattgccaaggtt 5032
|||||
Db 2286 AGGAACACCAGGAGGATCCTGAAATGTTGATTGATCTAATGTACAGAAATTGCCAAAGGTT 2345
QY 5033 accagacctctccaga--gcgattgacctggttgacagaacatgycaggcaagcactcag 5089
|||||
Db 2346 ACCAGACCTCTCCAGATCTGCGATTGACCTGGTTGCAGAACATGGCAAGCACTCAG 2405
QY 5090 aacgaagcaatcatgctgnaagctgacagtgcttagtccactcagcagcactgtgtctg 5149
|||||
Db 2406 AACGAAGCAATCATGCTGAAGCTGCACAGTGTCTAGTCCACTCAGCAGCACTTGTGCTG 2465
QY 5150 aatattlgagcatgctggaggagccggaatatcttccctgtgagtgtaacatttcaga 5209
|||||
Db 2466 AATATTGAGCATGCTGGAGAGACCGGAATATCTCTCTGTGGGATGTGTAACATTTCAGA 2525
QY 5210 atattcatctaatgttttagaagaatctgcygtctcagatgatgtgtatctccagatg 5269
|||||
Db 2526 ATATTTCATCTAATGTTTGAAGAATCTGCGCTCTCAGATGATGTGTATCTCCAGATG 2585
QY 5270 aagaagtlatctgctctggaataatcttactagtlcaggactgtgyggaattactggaac 5329
|||||
Db 2586 AAGAAAGTATCTGCTCGAAAAATACTTTACTGAGTCAGGACTGTGGGATTACTGGAAC 2645
QY 5330 aagcagctgctcctctctatagctgctgcatatgaagcagttaaatgaagttlacaag 5389
|||||
Db 2646 AAGCAGCTGCTCCCTCTCTATGCGCTGCACTATGAAGCAGTTAATGAAGTTTACAAAG 2705
QY 5390 tacttattccattcaatgaagctaatcgygtgcaaaagaactatccaaattcatgta 5449
|||||
Db 2706 TACTTATTCCTATTCATGAAGCTAATCGGGATGCAAGAATATCCACAATTCATGTGTA 2765
QY 5450 aacttcaagaagcatcagcaaaatgttctcatcagagtactgctggyagcgagatgttg 5509
|||||
Db 2766 AACTTCAAGAAGCATTCAGCAAAATGTTTCATCAGAGTACTGGCTGGGACCGGATGTTG 2825
QY 5510 gcacctatttctgtgtgttttltatggaaccaagltcgyggaattgagatgaaacaagaat 5569
|||||
Db 2826 GCACCTATTTTGGTGGTGTTTTATGGAACCAAGTTCGGGATTTGATGAACAAGAAAT 2885
QY 5570 ttgtttacaagagcctgcaataaaccaaaccttgcagagatatctcacagattgagagat 5629
|||||
Db 2886 TTGTTTACAAGAGCCCTGCAATAAACCAAACTTCAGAGATATCTCACAGATTGGAGGAT 2945
QY 5630 tttaaggagaaagattggagagagatgtgtgtgaagtaatcaagacttcaatcctgtag 5689
|||||
Db 2946 TTACGGAGAAAAAGATTGGAGAGGATGTGTTGAAGTAATCAAAAGACTTAATCCTGTAG 3005
QY 5690 acaagtgtlaaatlagatccctaacaagcatalattcagatttacctatgtgagccatact 5749
|||||
Db 3006 ACAAGTGTAAATTAGATTCCTAACAAAGGCATATATTACAGATTACCTATGTGAGCCATACT 3065
QY 5750 ttgacacatatgagatgaaggacagaatlcacctatttgcagaaaaaattacaatcttctgc 5809

Db 3066 TTGACACATATGAGATGAGAGACAGAAATCACCCTATTTCGACAAAAATTACAATCTTCGTC 3125
QY 5810 gattcatgtactgtacaccccttactttagatggcgttgcccatggygaaccttcatgaac 5869
Db 3126 GATTTCATGTACTGTACACCCCTTACTTTAGATGGCCGTGCCCATGGGGAACCTTCATGAAC 3185
QY 5870 aattcaaaagaagaccattctgactacgtcctcattgccttccctatatataaaacaagg 5929
Db 3186 AATTCAAAAGAAGACCACTTCTGACTACGCTTCATGCTTCCTTCTATATTAACAAGGG 3245
QY 5930 tcaatgtcactcataaagaagaagatcatccttaacaaccaatggaagtgctattgagaca 5989
Db 3246 TCAATGTCACTCATTAAGAAGAGATCATCTTAACACCAATTGAAGTGGCTATTGAGACA 3305
QY 5990 tgcagaaaaagaacacagaggttggcatttgcaacacatcagatcccgagaccccaaaa 6049
Db 3306 TGCAGAAAAAGACACAGAGAGTTGGCATTTGCACACATCAGGATCCCGACAGACCCCAAAA 3365
QY 6050 tgcttcagatgtgtactccaggatctgtaggaaccaacagtgaatcagggccttggaa 6109
Db 3366 TGCTTCAGATGGTACTCCAGGGATCTGTAGGCCACACAGTGAATCAGGGCCCTTTGGAAG 3425
QY 6110 ttgcccaagttttctgtctgaaataacctagtgaaccccaaggctctcagacatcataata 6169
Db 3426 TTGCCCAAGGTTTTCGTCTGAATACTTGTACCCCAAGCTCTTCAGACATCATATA 3485
QY 6170 aactgcgactcgtctttaaagatttactaaagggtgtgaagatgccttaagaaaaata 6229
Db 3486 AACTGCGACTCTGCTTTAAAGATTTTACTTAAAGGTGTGAAGATGCCTTTAAAGAAAAATA 3545
QY 6230 agagccttaattggccggttcaaaagaagatcaaaaggaattggggaactatctcgc 6289
Db 3546 AGAGCTTAATTGGCCGTTCAAAAGGAGTATCAAAAGGAATTGGGAAACTATCTTCGC 3605
QY 6290 cttaaagagccctacagccctagatcacagaaagtcctcagttatccaagccagttatt 6349
Db 3606 CTTAAAGAGGCCCTACAGCCCTAGATCACAGAAAGTCCCTCAGTATTCCAAGCCAGTATT 3665
QY 6350 gctgtccctgcacagaagattccctcagtcgaatgaagcttgcgaaaaatggtatctcta 6409
Db 3666 GCTTGCCCCCTGCCACAGAGATTCCTCAGTCGAATGAGCTTCGCAAAATGAGTCTCTA 3725
QY 6410 aactgaatgcactgttttatcatctgcgaagagccatgtattcaacatcgagtgtgaa 6469
Db 3726 AACTGAATGCACCTGTTTATTCATCTGCAGAGAGCCATGTATTCACATCGAGTGTGAA 3785
QY 6470 aagatctattggaacaaccaacatggaatgaaattctggaattaltattcaattgaagaatg 6529
Db 3786 AAGATCTATTGGAAACCAACATGGAATGGAATTCTGAAATTATTAATTCATGAAGAATG 3845
QY 6530 cagtgggccaagaaaaatcaaaatgtagatgtttaaagcttgagaatcattggtatgttt 6589
Db 3846 CAGTGCCCAAGAAATATCAAAATGTAGATTGTAACGCTTGAGAAATCATGGCTATGTTT 3905
QY 6590 ctaatgttcggttaacaagctgttatcctttaagacatttgaactcaaaagttacact 6649
Db 3906 CTAATGTTGGGTAAACAAGCTGTATCTTTAAAGACATTTTAATGACTCAAAAGGTACACT 3965
QY 6650 atacattaccattatataccattagctaaaggttaaaaaatttaccatttaagttcgt 6709
Db 3966 ATACATTTACCATTTATTTATACCATAGCTAAGGTTAAAAATTTATTCACCTTAAGTTCGT 4025
QY 6710 atttttaatttatatacaccatttatagattcatttggaccattttaaatgtagtaaat 6769
Db 4026 ATTTTAAATTTATATCACCATTTATAGATTCAATTTTGACCCCAATTTTAAATGTAGTAAT 4085
QY 6770 gcttattttaaaggtactaaaaaatatgtgaatgtttacctcgtgcgcgacaggcctc 6828
Db 4086 GCTTATTTTAAAGGTACTAAAAAATATGTGAATGTTTACCTCGTGGCGCCAGGGCCCTC 4144

RESULT 5

AX172980
LOCUS AX172980 4200 bp DNA linear PAT 03-JUL-2001
DEFINITION Sequence 107 from Patent WO0142297.
ACCESSION AX172980
VERSION AX172980.1 GI:14597958
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
1.4200
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="polynucleotide fragment"
BASE COUNT 1262 a 844 c 907 g 1187 t
ORIGIN

Query Match 60.4%; Score 4123; DB 6; Length 4200;
Best Local Similarity 99.9%; Pred. NO. 0;
Matches 4136; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

QY 2693 taaatcgtctctggaagccttagtaataagcaatccagatatatctggtaccccaagtcac 2752
Db 61 TAAATCGTTCTCGAAGCCCTAGTAATAGCAATCCAGATATATCTGGAGATCCCAAGTCAC 120
QY 2753 cagatgatgaagttcgtatcaatcatcgcggagtaaggtttagatcgtcccaattcctgg 2812
Db 121 CAGATGATGAAGTTCGATCAATCATCGGAGTAAGGTTTAGATCCGTTCAATTCCTGGG 180
QY 2813 ttaacactgtgtgtccaaaagctgccccatggygatccaaacccagtcgaagtgcagaat 2872
Db 181 TTAACACTGGTGTGCCAAAGACTGCCCATGGGATCCAAACCCAGTCCAAAGTGCAGAAAT 240
QY 2873 caacacaggtatgataagaaagttgtaatcgtatgtcttcgcacacagagacgtcaagtt 2932
Db 241 CAACACAGGCTATGATGCAAGTTGTAATCGTATGTCTTCGCACACAGACGTCGAAGTT 300
QY 2933 tcttaacaacatlaacggaacgcttaccactaaaaagcttltcacaggaagctggtc 2992
Db 301 TCTTACAACATTAAACGGGACGCTTACCACACTTAAAGCTTTTTCACGAGGAGCTGGCTT 360
QY 2993 tgcagtggtgttttgcaagtgycagcgcttcgggaatcagcttgcacaacagcctgttct 3052
Db 361 TGCAGTGGGTTGTTGGCAGTGGCAGCGTTCGGGAAATCAGCTTTGCAACAAGCCTGTTCT 420
QY 3053 ttttgaatlaatggttaagaagcatgtgtgcacatttatactttaatgataaacttgagg 3112
Db 421 TTTTGAATTAATGTTAAGAAGCATGTGTGCACCATTTTACTTTAATGATAAAGTTGAGG 480
QY 3113 ctccaaggaagaagtcgttttccagaacgcttcaatgatatgacattgcagctctgtcagca 3172
Db 481 CTCCAAGGAAAAAGTCGTTTCCAGAACGTTTCATGATGACATTTGCAGCTCTTGTCAACA 540
QY 3173 cgattgtcagtataagtttcaacgatttcagaagaagacacagaagaattgtagagactca 3232
Db 541 CGATTGCTAGTATATAGTTTTCACGATTTTCAGAAAGACACAGAAATGTTGAGAGACTCA 600
QY 3233 atacaagccttgatcttctcgaatgatcgtgtctgtttagagagagatttgttt 3292
Db 601 ATACAAGCCTTGCAATCTTCTCAATGATCTGTGTCTTATGAGACAGAGATTGTGTTT 660
QY 3293 ttagccttataaagtcctgctataaacaaggtgtcttcaaaagcttlaactattaccgaatc 3352
Db 661 TTAGCCTTATAAGTCTGCTATAAACAAGGTCTTCAAAAGCTTTACTCATTTACCGAATC 720
QY 3353 ccagtgctcgtgtccttgaggctggaatttctaagaatcatctgcagtcatgagcact 3412
Db 661 TTAGCCTTATAAGTCTGCTATAAACAAGGTCTTCAAAAGCTTTACTCATTTACCGAATC 720

Db 721 CCAGTGTCTGGTGTCCTTGAGGCTGGATTTCCTAGCAATCATCTGCAGTCATGAGCACT 780
QY 3413 atgttacattaaactlaccctgcagcttacttaccctgcatactccatcaactctcg 3472
Db 781 ATGTTACATTAACCTTACCCTGCAGCTTACTTACTCCACTGCATCTCCATCACTTCTCG 840
QY 3473 ttctctgcacaatctcagagttctgatttctcagaatgtlacaagaccaaaagattg 3532
Db 841 TTTCTCTGCACATCTCAGAGTTCTGGATTTCCTAGCAATGTACAAAGACCAAAAGATTG 900
QY 3533 caaatatgtttgaattatccggtgccttccgcacaacagcattatttgcagagactgtgt 3592
Db 901 CAAATATGTTGAATTAATCCGTCCTTCCGCCAACAGCATTAATTGGCAGGACTTGTGT 960
QY 3593 taacagagctgctgtcattttagaccctgatgctgaagagactgtttgattgcataaga 3652
Db 961 TTAACAGAGCTGGCTGTCAATTTAGACCCTGATGCTGAAGGACTGTTGGATTGCATTAAGA 1020
QY 3653 aagtcataatatagttacacaattactctccagtcacgactcagaacccgcgtactctg 3712
Db 1021 AAGTCATCAATATGTTAGTACACAATTTACTCTCCAGTCACGACTCAGACCCGGCTACTCTG 1080
QY 3713 accctcagataaaggctcgagtggccatgttgtatctacctctgattgttatcatg 3772
Db 1081 ACCCTCAGATAAAAGGCTCGAGTGGCCATGTTGATCTACCTCTGATTGGTATTCATG 1140
QY 3773 aaactgtacctcagctgtatgtatttacagaactcacaatcaacgaggaagaccaattt 3832
Db 1141 AAACGTACCTCAGCTGATGATTTTACAGAACTCACATCAACGAGAGACCAATTT 1200
QY 3833 gtatagccaactgatatgtatgaagaagtgaagcggaagtatgataagccagaccgtttgcca 3892
Db 1201 GTATAGCCACTGATGATTAATGAAGTGAGAGCGGAATGATTAAGCCAGACCGTTGCCA 1260
QY 3893 tggcaatcgcaaggacatcgttccctcaactaacaaggcctggcagtttccctcactcgt 3952
Db 1261 TGGCAATCGCAGGACATCGTCCCTCAACTAACAGGCTGGCAGTTTCCCTCCTCAGCT 1320
QY 3953 caacgagtggcaggcaacacactaccttllcaggcagaatcaagtlcgaaagcctttgatct 4012
Db 1321 CAACGAGTGGCAGGCAACACACTTCTTTCAGCAGAAATCAAGTGGAAAGCCTTTGATCT 1380
QY 4013 gtctacttgggttctcaaaaatgcagatgaacagttctacagaagtgtttacagatc 4072
Db 1381 GTCTACTTGGGTCTTCAAAAATGCAGATGAACAGTTCTACAGAAAGTGTTCAGATC 1440
QY 4073 tctcagttctgcagctaaacccgctatlagatctgcttlatctctgtgtcttctgttg 4132
Db 1441 TCTCAGTCTTGCACTAAACCCGGCTATTAGATCTGCTTATCTCTGTGTCTTGTCTTG 1500
QY 4133 agtataaagggaaaaaagtgtttgaaacgaatgaatagcttgacccttlaagaatacaaaag 4192
Db 1501 AGTATTAAGGGAATAAGTGTGTAACGAATGAATAGCTTGACCTTTAAGAAATCAAAAG 1560
QY 4193 acatgaagcaaaagcttgaagaagctatctctggagcataggtgcccagggcaagaatgg 4252
Db 1561 ACATGAGAGCAAAAGCTTGAAGAAAGCTATTCTTGGGAGCATAGGTGCCAGGCAAGAAATGG 1620
QY 4253 tacggcgaagccgaggaagcgtcgagagaagcccatctgaaagtgccttltgaaagtcaag 4312
Db 1621 TACGGCGAAGCCGAGGACAGCTCGAGAGAAAGCCCATCTGGAAGTGCCCTTTGGAAGTCAAG 1680
QY 4313 aaaatttgaggtggaagaaatatgactcactggtcgtcaaaacacagagaagcttgaca 4372
Db 1681 AAAATTGAGGTGAGGAAAGATATGACTCACTGCGCTCAAAACACAGAGAAGCTTGACA 1740
QY 4373 aatcaagagcagagattgaacacgaagcactgatttgatgtaaaacctgctacagaagcaa 4432
Db 1741 AATCAAGAGCAGAGATTGAACACGAACTGATGATGGAACCTGGCTACAGAGAGCA 1800
QY 4433 accataatctttagatacatatagagattgttgttcagaccgttctctgtaacggaatcca 4492
Db 1801 ACCTAATCATTTTAGATACATTAGAGATTGTTGTTCAGAGCCGTTTCTGTAAACGGAATCCA 1860

QY 4493 aagagagcatcttggtggaagtgtcaaaaagtgtctactacacagcatggcctgtaaccaaa 4552
Db 1861 AAGAGAGCATTTCTTGGTGAGGTGCTTAAAGTGCTACTACACAGCATGGCCTGTAAACCAA 1920
QY 4553 gtgcagtttatactacaacactgttttgcctacacagagagccttggttccaagtctctg 4612
Db 1921 GTGCAGTTTATCTACAACACTGTTTGGCTACACAGAGAGCCTTGGTTTCAAGTTTCTCG 1980
QY 4613 aactcttattgaagaagagacagagcagtgctgtgattttagtcctcaggtctccgac 4672
Db 1981 AACTCTTATTTGAAGAAGAGACAGAGAGAGTGTGCTGATTTATGCTCAGGCTTCTCCGAC 2040
QY 4673 actgtagcagtagcatcgttacatacgtlcaaccccagtgcccttactactaa 4732
Db 2041 ACTGTAGCAGTAGCATCGGTACAAATACGGTCAACCCCCAGTGCCCTTACTACTAA 2100
QY 4733 tgaagcaaaactttgagattgggaataacttggccaaggtttaaaatgcagtlaccaatgt 4792
Db 2101 TGAGGCAAAACTTTGAGATTGGGAATTACTTTGCCAGGGTTAAATGCAAGTACCATGT 2160
QY 4793 cactatcctccttggtggtgcacatctcagaattttaatgaagaattcttaagacgttctc 4852
Db 2161 CACTATCTCTCTTGGTGGGCACATCTTCAGAAATTTAATGAAGAAATCTTAAGACGTTCTC 2220
QY 4853 taaagaactatattgacatatgtctgaagaagatctggaattggaggaaacaacattctctg 4912
Db 2221 TAAAGACTATATTGACATATAGCTGAAGAAGATCTGGAATTGAGGGAAACAACATTTCTG 2280
QY 4913 atcaggtccagagatcgtgttccaatctccatatgattcttctgatactgtgaaatga 4972
Db 2281 ATCAGGTCAGGATCTGGTTTCAATCTCCATATGATTCCTTCGATACTGTGAATAATGA 2340
QY 4973 aggaacacagagagatcctgaaatgttgattgatactaatgtlacaagaattgccaaggtt 5032
Db 2341 AGGAACACACAGAGAGATCCTGAAATGTGATTGATCTAATGTACAGAAATTGCCAAGGTT 2400
QY 5033 accagacctctccaga--gcgattgacctggttcagaaacatggcaggcaagcactcag 5089
Db 2401 ACCAGACCTCTCCAGATCTGCGATTGACCTGGTTGCAGAACATGGCAAGCACCTCAG 2460
QY 5090 aacgaagcaatcatgtctgaagctgcacagtgcttagtccactcagcagcactgtgtctg 5149
Db 2461 AACGAAGCAATCATGCTGAAGCTGCACAGTGTCTAGTCCACTCAGCAGCACTTGTGCTG 2520
QY 5150 aatatltgagcatgctyggaagaccggaataatctcctgtggagatgtlaacattcaga 5209
Db 2521 AATATTGAGCATGCTGGAGGACCGGAATATCTTCTCTGTGGGATGTGAACATTTCAGA 2580
QY 5210 atattcatctaagtlttagaagaatctgcgtctcagatgagtglatctccagatg 5269
Db 2581 ATATTTCATCTAATGTTTTAGAAGAAATGCGGCTCTCAGATGATGTGGTATCTCCAGATG 2640
QY 5270 aagaagtatctgctctggaataacttactagtcagagactgtgggattactggaac 5329
Db 2641 AAGAAGTATCTGCTCTGGAATAATACTTACTGAGTCAGGACTGTGGGATTACTGGAAC 2700
QY 5330 aagcagctgcttccctctctatgctgcgtgatgtgaagcagttaatgaaglttacaag 5389
Db 2701 AAGCAGCTGCTTCCCTTCTATAGGCTGGCATGTATGAAGCAGTTAATGAAGTTTACAAG 2760
QY 5390 tacttattcctatcatgaagctaactcgggatgcaaaagaactatccaaatltcatgta 5449
Db 2761 TACTTATTCCTATTATGAAGCTAATCGGAGTGAAGAAGAACTATCCACAATTCAATGCTA 2820
QY 5450 aacttcaagaagcatcagcaaaatgttcatcagagtaactgctcggaggcggatgttg 5509
Db 2821 AACTTCAAGAAGCATTCAGCAAAATTTGTCATCAGAGTACTGGCTGGCAGCGGATGTTTG 2880
QY 5510 gcaacctatltcgtgtgtgttltatggaaccaagtlcgggagattggaatgaacaagaat 5569
Db 2881 GCACCTATTTTCGTGTGGTTTTATGGAACCAAGTTGCGGGATTTGGATGAACCAAGAAAT 2940

QY 5570 ttgtttacaaggagcctgcaataaaccaaactgcaagagatatctcacagaattgagagat 5629
|||||
Db 2941 TTGTTTACAAGGAGCCTGCATTAACCAAACTTCGAGAGATATCTCACAGATTTGGAGGAT 3000
QY 5630 tttaacggaagaatttgagagagatgtgtgtgaaglaatcaaaagacttaactctgtag 5689
|||||
Db 3001 TTACGGAGAAAGATTGGAGAGAGATGTGTTGAAGTAATCAAAAGACTTAATCCTGTAG 3060
QY 5690 acaagtgtaaatagatcctaacaagcatatatctcagattacctaigtgagccatact 5749
|||||
Db 3061 ACAAGTGAATATAGATTCCTAACAAAGCATATATTCAGATTACCTATGTGGAGCCATACT 3120
QY 5750 ttgacacatatgagatgaaggacagaatcacctatttcgacaaaaatracaatcttcgtc 5809
|||||
Db 3121 TTGACACATATGAGATGAAGAGACAGAAATCACCTATTTCGACAAAAATTACAATCTTCGTC 3180
QY 5810 gattcatgtactgtacaaccccttactttagatggccgtgcccataggggaaccttcataac 5869
|||||
Db 3181 GATTCACTGACTGTACACCCCTTACTTTAGATGGCCCGTCCCATGGGAACTTCATGAAC 3240
QY 5870 aattcaaaagaagaccatctcgaactacgtcatalgaccttccttataataaacaagg 5929
|||||
Db 3241 AATTCAAAAGGAAGACCATTTCTGACTACGTCATGCCCTTCCCTATATTAACAAGGG 3300
QY 5930 tcaatgtcactcataaagaagagatcatcttaacaccaatgaagtgtctattgagaca 5989
|||||
Db 3301 TCAATGTCACTCATAAAGAAGAGATCATCTTAACACCAATTGAAGTGTCTATTGAGACA 3360
QY 5990 tgcagaaaaagacacagaggagttgagcatttgcaacacatcagatcccgacagcccaaaa 6049
|||||
Db 3361 TGCAGAAAAAGACACAGGAGTTGGCATTTGCACACATCAGGATCCCGCAGACCCCAAAA 3420
QY 6050 tgcctcagatggtactccacagagatctgttagcaccacacagtgaaatcaggggaccttgaa 6109
|||||
Db 3421 TGCTTCAGATGGTACTCCAGGGATCTGTAGGCCACACAGTGAATCAGGGGCCCTTTGGAAG 3480
QY 6110 ttgcccaaggttttctgtctgaataacctagtgaccccaaaagctcttcagacatcataata 6169
|||||
Db 3481 TTGCCCAGGTTTCTGTCTGAATACTTGTGACCCAAAGCTCTTCAGACATCATATA 3540
QY 6170 aactgcgactcgtctttaaagatttactaaagggtgtgaagatgccttaagaaaaata 6229
|||||
Db 3541 AACTGCGACTCTGCTTTAAAGATTTTACTTAAAGGTGTGAAGATGCCCTTAAGAAAAATA 3600
QY 6230 agagcttaattgggcccgttcaaaaagafatacaaaaggaattggggaaactatcttcgc 6289
|||||
Db 3601 AGAGCTTAATTGGCCGTTCAAAAAGAGATATCAAAAGGAATTGGGGAAACTATCTTCGC 3660
QY 6290 cttaaagagccctacagccctagatcacagaaagtccctcagttatccaagccagtatt 6349
|||||
Db 3661 CTTAAAGAGGCCCTACAGCCCTAGATCACAGAAAGTCCCTCAGTTATCCCAAGCCAGTAT 3720
QY 6350 gcttgccccctgcacagagattccttcagtcgaatgagcttgcgcaaaatgagatctcta 6409
|||||
Db 3721 GCTTGTCCTGCCACAGAGATTCCTTCAGTCGAATGAGCTTTGCCAAATGGATCTCTTA 3780
QY 6410 aactgaatgcactgttttattcatctgcaaaagagccatgtattcaacaatcgagtgtgaa 6469
|||||
Db 3781 AACTGAATGCACCTTGTATTATTCATCTGCAAAAGAGCCATGTATTCACATCGAGTGTGAA 3840
QY 6470 aagatctatttgaaaccaacatggaatggaattctggaatttatattcatatgaagaatg 6529
|||||
Db 3841 AAGATCTATTGGAACCAACATGGAATGGAATTCTGGAATTAATTATTCATTGAAGAATG 3900
QY 6530 cagtggccaagaaaatatcaaatgtagatgttaacgcttgagaatcatgycatggttc 6589
|||||
Db 3901 CAGTGGCCAAGAAATATCAAAATGTAGATTGTTAACGCTTGAGAAATCATGGCTATGGT 3960
QY 6590 ctaatgttcgggtaacaagctgttatcttttaagacattttaatgactcaaaagtlacact 6649
|||||
Db 3961 CTAAATGTTCGGGTAAACAAGCTGTATCTTTTAAGACATTTTAATGACTCAAAAGGTACACT 4020
QY 6650 atacatttacattattataccatagctaaggttaaaaaatttacttaacttaagtgct 6709

Db 4021 ATACATTACCATTATTATTATACCATAGCTAAGGTTAAAAATTATTCACCTTAAGTTCGT 4080
QY 6710 atttttaatttatatcaccaatttatagattcattttggaaccattttaatgtagtaat 6769
|||||
Db 4081 ATTTTTAATTATATCACCAATTATAGATTCAATTTTGACCCCATTTTAAATGTAGTAAT 4140
QY 6770 gcttaatttaaaggtactaaaaaataatgtgaatgtttacacctgtgcgcgcgaaggccctc 6828
|||||
Db 4141 GCTTATTTTAAAGGTACTTAAAAAATATGTGAATGTTTACCTCGTGGCGCCAGGCCCTC 4199

RESULT 6
AX172978 3253 bp DNA linear PAT 03-JUL-2001
LOCUS AX172978
DEFINITION Sequence 105 from Patent WO0142297.
ACCESSION AX172978
VERSION AX172978.1 GI:14597957
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 3253)
AUTHORS Lu,P., Garman,J.D. and Candia,A.F.
TITLE Clasp-3 transmembrane protein
JOURNAL Patent: WO 0142297-A 105 14-JUN-2001;
Arbor Vita Corporation (US)
FEATURES
source location/Qualifiers
1..3253
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="polynucleotide fragment"

BASE COUNT 1009 a 633 c 705 g 906 t
ORIGIN

Query Match 47.3%; Score 3230.8; DB 6; Length 3253;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 3245; Conservative 0; Mismatches 2; Indels 3; Gaps 1;

QY 3582 aggaactgtgttaacagagctgctgtcattttagaccctgatgtcgaagagactgtttgg 3641
|||||
Db 3 AGGACTTGTTTACAGAGCTGGCTGTCAATTTTAGACCCGTGATGCTGAAGCACTGTTGG 62
QY 3642 attgcataaagaagtcatcaatatgtgtacacaatttactctccagtcgaagactcagacc 3701
|||||
Db 63 ATTGCATAAGAAAGTCATCAATATGTGTACACAAATTTACTCTCCAGTCACGACTCAGACCC 122
QY 3702 gcggtactctgaccctcaagataaaggctcgagtgccatgltgtatctacctctgattg 3761
|||||
Db 123 GCGGTACTCTGACCCTCAGATATAAGGCTCGAGTGGCCATGTTGTATCTACCTCTGATTGG 182
QY 3762 tattatcatggaactgtacactcagctgtatgattttacagaacctcaacaatcaacgag 3821
|||||
Db 183 TATTATCATGGAACCTGTACCTCAGCTGTATGATTTTACAGAAACTCAACATCAACGAGG 242
QY 3822 aagaccaattgtatagccactgtatgatgaagtgagagcggaagtatgataagcca 3881
|||||
Db 243 AAGACCAATTGTATAGCCACTGATGATTAAGTAGAGAGCGGAAGTATGATAAGCCA 302
QY 3882 gaccgttgccatgcgaatcgcagggacatcggtccctcaactaacaaggcctgycagttc 3941
|||||
Db 303 GACCGTTGCCATGGCAATCGCAGGACATCGGTCCCTCAACTAACAGGCCCTGGCAGTTT 362
QY 3942 cctccctacgtcaagagtgycagggcaacacactacccttcagcggaatcaagtgcgaag 4001
|||||
Db 363 CTTCTCACGTCAACGAGTGGCAGGCAACACACTACCTTTTCAGAGCAATCAAGTGAAG 422
QY 4002 ccttttgatctgtctactttggttctcaaaaatgcaatgaagaagttctacagaagt 4061
|||||
Db 423 CTTTGTATCTGTACTTTGGGTTCTCAAAAATGCAGATGAACAGTCTCTACAGAAGTG 482
QY 4062 gtttacagatctctcagctcttgcaagctaaacggctatttagatctgcttattctctgt 4121

|||||
Db 483 GTTACAGATCTCTCAGTCTTGACGCTAAACCGGCTATTAGATCTGCTTATCTCTGTGT 542
QY 4122 gtcttgctttgagtataaaggaaaaagtggtttgaaagaaatgaatagcttgacctttaa 4181
Db 543 GTCTGCTTTGAGTATTAAGCGAAAAAGTGTGTGAACCAATGAATAGCTTGACCTTTAA 602
QY 4182 gaaatcaaaaagacatgagagcaaaagcttgaagaagctattcttggagcatagtgccag 4241
Db 603 GAAATCAAAAAGACATGAGAGCAAAAGCTTGAAAGAACTATTCTTGAGCATAGTGCCAG 662
QY 4242 gcaagaaatggtacgycgaagccgagagacagctcgagagaaagcccatctgnaagtgccct 4301
Db 663 GCAAGAAATGGTACGGCGAAGCGAGACAGCTCGAGAGAAAGCCCATCTGAAAGTGCCCTT 722
QY 4302 tggaaagtcaagaaatattgaggtggaggaagatatgactcactgycgtcaaaacacaga 4361
Db 723 TGGAAGTCAAGAAAATTGAGGTGGAGGAAAGATATGACTCACTGGCGTCAAAACACAGA 782
QY 4362 gaagcttgacaatatcaagagcagagattgaaacacgaagcaactgattgatgnaaactgyc 4421
Db 783 GAAGCTTGACCAATCAAGAGCAGAGATTGAACACGAAGCACTGATTGATGGAACCTGGC 842
QY 4422 tacagaagcaaacctaatactttagatatacattagagattggttcagaccgttctgt 4481
Db 843 TACAGAGCAAAACCTAATCATTTTAGATACATTAGAGATTGTTGTTCAGACCCGTTTCTGT 902
QY 4482 aacggaatccaaaagagagacattcttgtygagtgctaaaagtgctactacacagatgyc 4541
Db 903 AACGGAATCCAAAAGAGAGACATCTTGTTGGTGGAGTGCTAAAGTGCTACTACACAGCATGGC 962
QY 4542 ctgtaaccaaaagtgacgtttatctacaacactggtttgtctacacagagagccttgcttc 4601
Db 963 CTGTAAACCAAAAGTGAGTTATCTTACAAACACTGTTTGGTACACAGAGAGCCTTGTTTC 1022
QY 4602 aaagttccctgaactcttatttgaagaagagacagagcagtgtygtatttatgcttcag 4661
Db 1023 AAAGTTTCCTGAACCTCTATTGTGAAGAAGAGACAGACAGCAGTGTGATTTATGCCTCAG 1082
QY 4662 gcttctccgacactgttagcagtagcatcggtacaataacggtcaacaccccaagtgcctcc 4721
Db 1083 GCTTCTCCGACACTGTAGCAGTAGCATCGGTACATACGGTCAACACCCAGTGCCTCCCT 1142
QY 4722 ttacctactaatgagggcaaaactttgagattgggaataactttgccagggttaaaatgca 4781
Db 1143 TTACCTACTAATGAGGCAAAACTTTGAGATTGGGAATACTTTGCCAGGGTTAAATGCA 1202
QY 4782 ggtaccaatgtcactatacctccttgtytggycacatctcagaaatttlaatgaagaattctt 4841
Db 1203 GGTAACCAATGTCACTATCTCTTGGTGGGCACATCTCAGAAATTTTAATGAAGAATTTCT 1262
QY 4842 aagacgttctctaaagactataatgacataatgctgaagaagatctgnaattgagggaaac 4901
Db 1263 AAGACGTTCTCTAAAGACTATATGTGACATATGCTGAAGAAGATCTCTGAATTGAGGGAAC 1322
QY 4902 aacatttccgtatcaggtcccaagatctggttttcaatctccatatgatatcttcttgatc 4961
Db 1323 AACATTTTCCTGATCAGGTCCAGGATCTGGTTTTCATATCTCCATATGATTCTTCTGATAC 1382
QY 4962 tgtgaaaatgaaggaacacccagggagatccctgaaatgttgattgattctaatgtacagaat 5021
Db 1383 TGTGAAAATGAAGGAACACCGAGAGATCCTGAATGTGATTGATCTAATGTACAGAAAT 1442
QY 5022 tgcacaaggttaccaagactctccaga---gcgattgacctggtttgcagaacatgycag 5078
Db 1443 TGCCAAAGGTTTACGAGACCTCTCCAGATCTGCGATTGACCTGGTTGCAGAACATGGCAGG 1502
QY 5079 caagcaactcaagaacgaagaatcagtctgaagctgcacagtgcttagtccactcaagcagc 5138
Db 1503 CAAGCACTCAGAAACGAACATCATCTGAAAGCTGCACAGTGTCTAGTCCACTCAGCAGC 1562
QY 5139 acttgcttgtaatatattgagcatgctggaagaccggaatatcttccgtgtygagtgt 5198
|||||

Db 1563 ACTTGTTGCTGAATATTTGAGCATGCTTGAGGAGCACCAGAAATATCTTCTGTGGGATGTGT 1622
QY 5199 aacatttcagaatatattcatctaaatgltttagaagaatctgcggtctcogatgtatgtgt 5258
Db 1623 AACATTTCAGATATTTTCATCTAATGTTTTAGAAGAATCTGGCGTCTCAGATGATGTGT 1682
QY 5259 atctccagatgaagaaggtatctgctctcggaaaatactttactgagtcagagactgtgtgg 5318
Db 1683 ATCTCCAGATGAAGAAGGTATCTGCTCGGAAAATACTTTACTGAGTCAGACTGTGGG 1742
QY 5319 attactggaacaagcagctgcttccctctctatgctgtygcagtatatgaagcaqtltaatga 5378
Db 1743 ATTACTGGAACAAGCAGCTGCTTCTCTATGGCTGGCATGTATGAAGCAGTTAATGA 1802
QY 5379 agttacaagtaacttattccatattcaatgaagctaactcggaatgcaagaactatccac 5438
Db 1803 AGTTTACAAGTACTTATTCTTATTCATGAAGCTAATCGGGATGCAAGAACATAITCCAC 1862
QY 5439 aattcatgtaaaactlcaagaagcatlcaagcaaaaatgltcatcagagtlactgctggga 5498
Db 1863 AATTCATGTTAACTTCAAGAAGCATTCAGCAAAATTGTTTCATCAGAGTACTGGCTGGGA 1922
QY 5499 gcggaatgttggcacctatttctgtygtgttltatgnaaccaagtlcgggatttga 5558
Db 1923 GCGGATGTTGGCACTATTTTCGTGTGGTTTTATGGAACCAAGTTCGGGATTTGGA 1982
QY 5559 tgaacaagaatttgtttacaaggaagcctgcaataaccaacttgcagagatatctcacag 5618
Db 1983 TGAACAGAAATTTGTTTACAGAGAGCCGTCGAATAACCAACTGCGAGATATCTCACAG 2042
QY 5619 attggaaggattttacggygaagaagatttggagagatgtgtgtgaagtaatcaaaagactc 5678
Db 2043 ATTGAAGGATTTTACGGAAGAAAGATTGGAGAGGATGTGTTGAAGTAATCAAAAGACTC 2102
QY 5679 taatcctgtagacaagtgtaaattagatccctaaacaagycatatattcagattactatgt 5738
Db 2103 TAATCTGTAGACAGTGTAAATTAGATCCTTAACAAGGCATATATTCAGATTACCTATGT 2162
QY 5739 ggaaccatacttgaacaacatatgagatgaaggaacagaatcacctaattcgacaaaatta 5798
Db 2163 GGAGCCATACTTGAACACATATGAGATGAAGACAGAAATCACCTATTTTCACAAAAATTA 2222
QY 5799 caatcttgcgtcatlcatgtlactgtacaccccttactttagatgycggtgccatggyga 5858
Db 2223 CAATCTCTGTCGATTCATGTACTGTACACCTTTACTTTAGATGGCCGTGCCCATGGGA 2282
QY 5859 actlcatgaacaattcaaaaagaagaccaattctgactacgtctcaatgaccttccctatat 5918
Db 2283 ACTTCATGAACAATTCAAAAAGAGACCAATTCGACTACGTCTCATGCTTCTCTATAT 2342
QY 5919 taaaacaaggytcaatgtcactcatataaagaagagatcatcttaacaccaatttgaagtgc 5978
Db 2343 TAAACCAAGGGTCAATGTCACTCATTAAGAAGAGATCATCTTAACACCAATTGGAAGTTGC 2402
QY 5979 tattgagacatgcagaaaaaagacacaggaagttggcatttgcacaacatcagatccgcg 6038
Db 2403 TATTGAGGACATGCAGAAAAAGACACAGAGTGGCATTTTGCAACACATCAGGATCCCGC 2462
QY 6039 agaccccaaatgctcagatggtlactccaggaatctgtagccacacagtgaatcaggg 6098
Db 2463 AGACCCCAAAATGCTTCAGATGGTACTCCAGGGATCTGTAGGCACACAGTGAATCAGGG 2522
QY 6099 gccttggaaagttgcccaaggttttctgtctgaaatacactagtgacccaagctctcag 6158
Db 2523 GCCTTTGGAAGTTGCCCAGGTTTTCGTCTGAAATACCTAGTGACCCAAAAGCTCTTCAG 2582
QY 6159 acatcataataaactgagactctgctttaagatttactataaaggttgaagatgcctt 6218
Db 2583 ACATCATATAATACTGCGACTCTGCTTTAAAGATTTTACTAAAGAGTGTGAAGATGCCTT 2642
QY 6219 aagaaaaataagagctttaatltgggcccgttcaaaaagagatatcaaaaggaattggygaa 6278
Db 2643 AAGAAAAAATAAGAGCTTAATTGGGCCGCTTCAAAAGAGATATCAAAAGGAATTGGGGA 2702
|||||

QY	6279	actatcttcgccttaagagggccctacagcccttagatcacagaagaagtcctcagttatcc	6338
Db	2703		
QY	6339	aagccagtatctgtctgtccctgcacagagatccttcagtcgaatgagctttcgcaa	6398
Db	2763		
QY	6399	atggatctctaaactgaatgacactgttttattcatctgcaaaagaccatgtattcaaca	6458
Db	2823		
QY	6459	tccagtgtaaaagatcatttgaaaccaacatggaatctcgaattattatctc	6518
Db	2883		
QY	6519	attgaagaatgcagtggcacaaagaatatcaaatgtagattgtttaacgcttgagaatcat	6578
Db	2943		
QY	6579	ggctatggttctctaattcgcgggtaacaagctgttatcttttaagacattttaatgactc	6638
Db	3003		
QY	6639	aaaggtacactatacattaccattattatcacatagctaaggtlaaaaaattatcac	6698
Db	3063		
QY	6699	tttaagtcgtaatttttaatttatatcacattatagattcatttgagaccatttta	6758
Db	3123		
QY	6759	aatgtagtaatgcttattttaaaggtactaaaaataatgtgaatgttaccctcgtcgcg	6818
Db	3183		
QY	6819	ccagggcctc	6828
Db	3243		
RESULT	7		
AK055905		3327 bp	mRNA linear PRI 31-OCT-2001
LOCUS		Homo sapiens cDNA FLJ131343 fls, clone MESAN1000101, weakly similar	
DEFINITION		to Rat try gene product.	
ACCESSION	AK055905		
VERSION	AK055905.1	GI:16550749	
KEYWORDS		oligo capping; fls (full insert sequence).	
SOURCE		Homo sapiens normal mesangial cells (NHMC56046-2) cDNA to mRNA, clone lib:MESAN1 clone:MESAN1000101.	
ORGANISM		Homo sapiens	
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
AUTHORS		1 (sites)	
		Tashiro,H., Yamazaki,M., Watanabe,K., Kumagai,A., Itakura,S., Fukuzumi,Y., Fujimori,Y., Komiyama,M., Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H., Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K., Wagatsuma,M., Murakawa,K., Kanehori,K., Takahashi-Fujii,A., Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y., Sugano,S., Nagahari,K., Masuho,Y., Nagai,K. and Isogai,T.	
TITLE		NEDO human cDNA sequencing project	
JOURNAL		Unpublished	
REFERENCE		2 (bases 1 to 3327)	
AUTHORS		Isogai,T., Otsuki,T. and Sugiyama,T.	
TITLE		Direct Submission	
JOURNAL		Submitted (24-OCT-2001) Takao Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:genomics@hri.co.jp, Tel:81-438-52-3951, Fax:81-438-52-3952) NEDO human cDNA sequencing project supported by Ministry of	
COMMENT			

Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: RAB and HRI.

FEATURES
source

Location/Qualifiers
1..3327
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MESAN1000101"
/cell_type="normal mesangial cells (NHMC56046-2)"
/note="cloning vector: pME18SFL3~primary culture, normal mesangial cells"
54..2279
/note="unnamed protein product"
/codon_start=1
/protein_id="BAB71042.1"
/db_xref="GI:16550750"

CDS

translation="MFELSVPERQOHYLAGLVLTETLAVILDDPAEGLFGLHKKVINM
HNLLSHSDPDRYSDPQIKARVAMLYLPLIGIMETVPOLYDPTETHNQGRPICAT
DDYESESGSMISQIVAMIAIGTSVPQLRPGSFLSTSGRHTTFESAESSRLICL
LWLNKNADEIVLQKWFDTLSVLQNLRLDLLYLCVSCFEYKQKVFERNMSLTPEKSK
DMRAKLEELIGSIGARQEMVRSRQDERSPSGSAFQSEQENRWRKDMTHWRQNTK
LDKSRAEIEHALLIDGNLATEANLIIDLEIVQTVSVTESKESILGVLKVLHSM
ACNOSAVYLQHCFFATQRLALVSKFPELLFEETEOCADLCRLRHCCSSIGTIRSHAS
ASLYLLMRQNEIIGNNFARVKMQVMTSLSLVGTSONFNEFLRRSLKILTYAEDL
ELRETFPQDOVDLVENLHMLISDQVYKMKHEQEDPEMLIDLMYRIAKGYQTSPLRLT
WLQNMAGKHSBERSNHAEOCLVHSAALVAEYLSMLEDRKYLPGCVTFONISSNLE
ESAVSDVYSPDEEGICSGKYFTESGLVGLLEQAAFSMAGMYEAVNEVYKVLPIH
EANRDAKLSTIHGKLQEAFSKIVHQDKRMFGTYFRVGFQGTKEFGDLDEQEFVYKEP
AITKLAEISHRLEGFYGERFGEVDVVEIIKDSNPVDCKLDLPNTHIFRLPMMSHTLTHM
R"

BASE COUNT 1046 a 640 c 711 g 930 t
ORIGIN

Query Match 46.9%; Score 3201.8; DB 9; Length 3327;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 3285; Conservative 0; Mismatches 32; Indels 14; Gaps 5;

QY	3485	catctcagagttctgattcttctcagaatgtlacaagcaaaagattgcaaatatgttg	3544
Db	1	CATCTCAGAGTTCTTGATTTCTTACGAATGTACAGACCACAAAGATTGCAATATGTTTG	60
QY	3545	aattaccgtgccttccgcacaaagcattatttgcaggaactgtgttaacagagctg	3604
Db	61	AATTATCCGTGCGCTTCCGCCAACAGCATTTATTGGCAGGACTGTGTTAACAGAGCTGG	120
QY	3605	ctgtcattttagaccctgtagctgtaagagactgtttggaattgcatagaagaatcatcata	3664
Db	121	CTGTCAATTTTAGACCCCTGATGCTGAAGGACTGTTTGCAATGAAGACTCATCAATA	180
QY	3665	tggtacacaaattactctccagtlcaagctcagaccgcggtactctgaccctcagataa	3724
Db	181	TGGTACACAATTTACTCTCCAGTCAAGACTCAGACCGCGGTACTCTGACCTCAGATAA	240
QY	3725	aggctcagtggtccatgtgtatctacactctgattggtatlatcatgnaactgttacctc	3784
Db	241	AGGCTCGAGTGGCCATGTTGATCTTACCTCTGATTCGTATTATCATGGAACCTGTACCTC	300
QY	3785	agctgtatgatltttacagaactcacatcaacgaggaagaccaattgttatagccactg	3844
Db	301	AGCTGTATGATTTTACAGAAACTCACAAATCAACAGAGGAAGACCAATTTGTATAGCCACTG	360
QY	3845	atgattatgaagtgagagcggaagtatgatgaagccagaccgtttgccatgcaatcgag	3904
Db	361	ATGATTTATGAAGTGAGAGCGGAAGTATGATGATAGCCAGACCGGTGCCATTCGCAG	420
QY	3905	ggacatcggtccctcaactaacaaggcctgagcttctcctcacgltcaacgagtgga	3964

Db 421 GGACATCGGTCCCTCAACTAACAAAGGCCCTGGCAGTTTCCCTCCTCAAGCTCAACGAGTGGCA 480

QY 3965 ggcacaacactcccttlttcagcagaatacgaagtcgaagcccttltgaatctgtctactttgg 4024

Db 481 GGCAACACACTACTCTTTTCAGCAGAATCAAGTGAAGCCCTTTGATCTGTCTACTTTGGG 540

QY 4025 ttctcaaaaatgcagatgaacacagttctacagaagtggttacagatctctcagtcctgc 4084

Db 541 TTCTCAAAAATGCAGATGAACAGTTCTTACAGAAAGTGTTACAGATCTCTCAGTCTTGC 600

QY 4085 agctaaaccggtattagatcgtcttactctctgtgtgtctgtctgtctgttgagtataaggga 4144

Db 601 AGCTAAACCGGCTATTAGATCTGCTTATCTCTGTGTGTCTTGCCTTGAGTATAAAGGA 660

QY 4145 aaaaagtgttgaacggaatgaatagcttgacccttlaagaaatcaaaaagacatgagagcaa 4204

Db 661 AAAAAGTGTGGAACGAATGAATAGCTTGACCTTTAAGAAATCAAAAAGACATGAGAGCAA 720

QY 4205 agcttgaagaagctattctcttggagcataagtgccagggcaagaaatggtacgycgaagcc 4264

Db 721 AGCTGAAGAGAGCTATTCTTGGAGCATAGGTGCCAGGCAAGAAATGGTACGGCGAAGCC 780

QY 4265 gaggaacagctcgaagaagccccatctggaagtgcccttgaagtcagaagaaattgaagt 4324

Db 781 GAGGACAGCTCGAAGAGAACCCCATCTGGAAGTGCCCTTTGGAAGTCAAGAAATTTGAGGT 840

QY 4325 ggaagaaagatatgaactcactgycgtccaaaaacacagagaagcttgacaaatcaagagcag 4384

Db 841 GGAGAAAGATATGACTCACTGGCGCTCAAAAACACAGAGAAGCTTGACAAATCAAGAGCAG 900

QY 4385 agattgaacagaagcactgaattgtagtgaaacctggtctacagaagcaaacctaattc 4444

Db 901 AGATTGAACACGAAGCAGCACTGATGTGAGAAACCTGGCTACAGAACCAACCTATCATTT 960

QY 4445 tagatatactagagattgttlttcagaccgttctgttaacggaatccaaagagacatcc 4504

Db 961 TAGATACATFAGAGATTGTGTTCAGACCGTTTCTGTACCGGAATCCAAAGAGAGCATTC 1020

QY 4505 ttggtgagtgctaaagtgtcactacacagcatgycctgttaaccaaagtgcaagttatc 4564

Db 1021 TTGTTGAGTGTCTAAAGTGTCTACTACACAGACCATGGCCGTGAACCAAGTGCAAGTTATTC 1080

QY 4565 tacaacactgttlttgcctacacagaagcccttggttccaagttccctgaactcttatttg 4624

Db 1081 TACAACACTGTTTTTGTCTACACAGAGACCTTGTTTCAAAAGTTTCTCTGAACCTCTTATTTG 1140

QY 4625 aagaagagacagaagcagtggtgtgattatgcctcagagcttccgaacactgtagcagta 4684

Db 1141 AAGAAGAGACAGAGCAGTGTGCTGATTTATGCTCAGGCTTCTCCGACACTGTAGCAGTA 1200

QY 4685 gcatcgtaacaatacgttcacaccccagtgcccttacccttaactactaatgagcgaacct 4744

Db 1201 GCATCGGTACAATAACGGTCACACGCCAGTGCCCTCCCTTACTACTAATGAGGCCAAACT 1260

QY 4745 ttgagattgggaataactttgcccaggtttaaattgcaggtaccaatgtcactatccctc 4804

Db 1261 TTGAGATTGGGAATAACTTTGGCAGGGTTAAATGCAGGTAAACAATGTCATATCCTCTC 1320

QY 4805 tgytggcacatactcagaattttaaagaagaattccttaagaagttctcctaaagactatat 4864

Db 1321 TGGTGGGCACATCTCAGAATTTTATGAAGAATTCTTAAGACGTCTCTTAAGACTATAT 1380

QY 4865 tgacatatctgaagaagactcgtgaattgaagggaacaacacattcctgatcaggtccagg 4924

Db 1381 TGACATATGCTGAAGAAGATCTGGAATTGAAGGAAACAACATTTCTGATCAGGTCAGG 1440

QY 4925 atctggttccaatctccatatgatcttctgatactgtgaaaaatgaagaagcaccagg 4984

Db 1441 ATCTGTTTCAATCTCATATGATTTCTTCTGATACTGTGAATAAGGAACACACAGG 1500

QY 4985 aggatcctgaaatgttgattgatcctaattgtacagaattgccaaggttacagacctctc 5044

Db 1501 AGGATCCTGAATGTTGATTGATCTAATGTACAGAATTGCCAAGGGTTACCAGACCTCTC 1560

QY 5045 caga---gcgattgacctggttgcagaacatggcgaagcagcactcagaacggaacatc 5101

Db 1561 CAGATCTGCAGCTGACTGCTGGTTGCAGAACATGGCAGGCAAGCAGCACTCAGAACGAACATC 1620

QY 5102 atgctgaagctgcacagtgctcagtcacctcagcagcactgttgcgtgaatatltgagca 5161

Db 1621 ATGCTGAAGCTGCACAGTGTCTAGTCCACATCAGCAGCACTTGTGCTGAATATTTGAGCA 1680

QY 5162 tgcctgagcaccggaatatctccctgtgtggaatgylaacattcagaatatltcatcta 5221

Db 1681 TGCTGAGGAGCCGGAATATCTTCCCTGTGGGATGTGTAAACATTTCAGATATTTCACTCTA 1740

QY 5222 atgtttagaagaatctgcggtctcagatgagtgtatctccagatgaagaaggtatct 5281

Db 1741 ATGTTTGAAGAATCTGCCGCTCAGATGATGGTATCTCCAGATGAAGAAGTATCT 1800

QY 5282 gctctggaataacttactagtcagagactgttggatltactggaacaagcagctgct 5341

Db 1801 GCTCTGGAATAATACTTTACTGAGTCAGACTGTGGGATTACTGGAACAAGCAGCTGCTT 1860

QY 5342 ccttctctatgctgcatgtlatgaagcagttaatgaagtttacaagtacttctcta 5401

Db 1861 CTTCTCTATGGCTGGCATGTATGAAGCAGTTAATGAAGTTTACAAGTACTTATTCCTA 1920

QY 5402 ttcatgaagctaatcgggaltgcaaagaacataccaaatltcatgtgaacttcaagaag 5461

Db 1921 TTCATGAAGCTAATCGGAGTGCAAAAGAACTATCCAAATTCATGTGTTAAACTTCAAGAAG 1980

QY 5462 caltcagcaaatgttcatcagagtagtgcgtggagcggatgttggcaacctatttc 5521

Db 1981 CATTCAAGCAAAATTTGTCATCAG-----GATGGTAAAGCGGATGTTGGCACCTATTTTC 2034

QY 5522 gtgtgtgttltatgnaaccaaagttcggggaatttgatgaacaaagaattgtttacaag 5581

Db 2035 GTGTGGTTTTTATGGAACCAAGTTCGGGGAATTGGATGAACAAGAATTGTTACAAAG 2094

QY 5582 agcctgcaataaccaaaacttgcagagatatctcacagatttgaagggaatttcaagga 5641

Db 2095 AGCCTGCAATTAACCAAACTTGCAGAGATATCTCACAGATTGGAGGGATTTTACGGAGAA 2154

QY 5642 gatttggagagatgtgttgaagtaatcaaaagctctaactcctgtagaacaagtgtaat 5701

Db 2155 GATTTGAGAGGATGTGTTGAAGTAATCAAGACTCTAATCCTGTAGACAAGTGAAT 2214

QY 5702 tagatcctaacaagcatatatltcagattacctatgttgaagccatacttgaacacatatg 5761

Db 2215 TAGATCCTAAC-ACGCATATATTCAGATTACTATGTGAGGCCATCTTTGACACATATG 2273

QY 5762 agatgaagcagaatacacctatttcgacaaaaatlacaactctgcgtcatcgttact 5821

Db 2274 AGATGAAGGACAGAAATCAACCTATTTGCACAAAATTACAATCTTCGATTCATGTACT 2333

QY 5822 gtacacccttactttagatgycggtgcccattggggaacttcatgaacaattccaaga 5881

Db 2334 GTACACCCCTTACTTTAGATGGCCGTGCCCATGGGGAACCTTCATGAACAATCAAAAGA 2393

QY 5882 agaccattctgactacgtctcatgyccttcccttatattlaaaacaaggttcaatgtcactc 5941

Db 2394 AGACCATCTGACTACGTCTCATGTGCCCTTTCCTTATATTAAACAAGGSTCAATGTCACTC 2453

QY 5942 ataaagaagagatcatcttacaaccaatlyaagltgctatgtgaagacatgcaagaaaaa 6001

Db 2454 ATAAAGAAGAGATCATCTTAACACCAAATGAAGTTGCTATTGAGGACATGCAGAAAAAGA 2513

QY 6002 caaaggaattggaatttgcacaacatcagaatcccgagagccccaanaatgcttcagatg 6061

Db 2514 CACAGGAGTTGGCAATTGCAACACATCAGGATCCCGCAGACCCCAAAATGCTTCAGATGG 2573

QY 6062 tactccaggaatctgtaggcaccacagltgaatcaaggyccttggaaagttgccaggttc 6121

Db 2574 TACTCCAGGATCTGTAGGCACACACAGTGAATCAGGGCCTTTGGAAAGTTGCCCAGGTTT 2633

QY 6122 ttctgtctgaaatacctagtgaccccaagctcttcagacatcataataaactgagctct 6181
|||||
Db 2634 TTCTGTCTGAAATACCTAGTGACCCCAAGCTCTTCAGACATCATATAACTGGCAGCTCT 2693
QY 6182 gctttaagaatttacttaaaaggtgtgaagatgccttaagaaaaataagagcttaatg 6241
|||||
Db 2694 GCTTTAAAGATTACTTAAAGGTGTGAGATGCCCTTAAGAAAAATAAGAGCTTAATG 2753
QY 6242 ggcgggttcaaaaggagltcaaaaggaaatg--gggaaactatctcgcccttaagagg 6299
|||||
Db 2754 GGCCGGATCAAAAGAGATCAAAAGGGAAGTGGAGAGAAACTATCATCGCCTTAAAGAGG 2813
QY 6300 ccctacagccctagatca--cagaaagtcctcagttatccaagccagtatgtgtcc 6357
|||||
Db 2814 CCCTACAGCCACTGATCAACAGAAAGATCCCTCAGTTATACAAGCAGTATTGCCGTGCA 2873
QY 6358 cctgccaagagatccttcagtcgaatgagcttcgcaaatgagatctctaactgaat 6417
|||||
Db 2874 CCTGCCACAGAGATTCCCTCAGTCGAATGAGCCTTCGCAAAATGGATCTTAACCTGAAT 2933
QY 6418 gcaactgttttatcatctgcгааaggccatgtattcaacatcgagtgtaaagatcta 6477
|||||
Db 2934 GCACCTGTTTATTCATCTGCAAGAGCCATGTATTCACATCGAGTGTGAAGAAGATCTA 2993
QY 6478 ttggaaccaacatggaatggaatctctgaaattattatcattgaagaatgcaagtgcc 6537
|||||
Db 2994 TTGGAACAACACATGGAATGGAATTCTGGAATTTATTATTCATGGAAGATGCAGTGGCC 3053
QY 6538 aagaaataatcaaatgtagattgtttaacgcttgagaatcatggtcgtatcttaatgtt 6597
|||||
Db 3054 AAGAAATATCAAAATGTAGATTGTTAACGCTTGAGAAATCATGGCTATGGTTCTAATGTT 3113
QY 6598 cgggtacaacagtgtatcctttaagacatttaactcaaggtacaactatacatt 6657
|||||
Db 3114 CTGTTAACAAAGCTGTTATCTTTTAAGACATTTTAATGACTCAAAAGGTACACTATACATTT 3173
QY 6658 accattattatacactagctaaggttaaaattattcactttaagtcgtattttta 6717
|||||
Db 3174 ACCATTATTATACCATAGCTAAGGTTAAAAATTATTTCACCTTAACTTCGATTTTT 3233
QY 6718 attatataccattatagattcatttttgaccattttaaatgtagtaatgctattt 6777
|||||
Db 3234 ATTTATATTACCATTTATAGATTTCATTTTGAACCATTTTAAATGTAGTAAATGCTTATTT 3293
QY 6778 taaagttactaaaaaatatgtgaatgtttac 6808
|||||
Db 3294 TAAAGGTACTATTAAATATGTGAATGTTTAC 3324
RESULT 8
AK055493 2452 bp mRNA linear PRI 31-OCT-2001
LOCUS Homo sapiens cDNA FLJ30931 fis, clone FEBRA2006873, weakly similar
DEFINITION to Trg protein.
ACCESSION AK055493
VERSION AK055493.1 GI:16550228
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens fetus brain cDNA to mRNA, clone_lib:FEBRA2
clone:FEBRA2006873.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS 1 (sites)
Nishi,T., Nakagawa,S., Senoh,A., Mizuguchi,H., Inagaki,H.,
Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S.,
Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T.,
Kimura,K., Yamashita,H., Matsuo,K., Nakamura,Y., Sekine,M.,
Kikuchi,H., Kanda,K., Wagatsuma,M., Murakawa,K., Kanehori,K.,
Takahashi-Fujii,A., Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y.,
Sugano,S., Nagahari,K., Masuno,Y., Nagai,K. and Isogai,T.
TITLE NEDO human cDNA sequencing project
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2452)

AUTHORS Isogai,T., Otsuki,T. and Sugiyama,T.
TITLE Direct Submission
JOURNAL Submitted (24-OCT-2001) Takao Isogai, Helix Research Institute,
Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3951, Fax:81-438-52-3952)
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction: Helix Research Institute (HRI) (supported by Japan
key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: RAB and
HRI.
FEATURES
source Location/Qualifiers
1. 2452
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="FEBRA2006873"
/tissue_type="brain"
/clone_lib="FEBRA2"
/dev_stage="fetus"
/note="Cloning vector: pME18SFL3"
171. 2051
/note="unnamed protein product"
/codon_start=1
/protein_id="BAB70933.1"
/db_xref="GI:16550229"
/translation="MACNONAVYLQHCFAIORALVSKPELLFEEETEEOCADICRLRL
RHCSSIGTIRSHASASLYLKRONEFIGNFARVKMOYTWLSLVLGTSQNFNEFL
RRSLKTLITYAEEDLEIRETFEPDOVDLVFNLMHLSITVKMKEHOEDPEMLIDIMY
RIAKGYQTSDDLRLTWLQNMAGKHSERSNHAEAQCLVHSALVAEYLSMLEDKYLP
VGCVFQNTSSNVLSESAVDVSPDEIGCSGKVFESGLVGLLEOAAASFNAGM
YEAVNEYKVLIPiHEANRDARKLSTIHGKLQEAFSKIYHOSGTWERMFGTYFRVGFY
GTFEGDLEQKVFYKEPATIKLAELSHRLEGFYGERFGDEVVEVIKDSNPVCKLDP
NKAYIQITYVEPYFDYEMKDRITFDKNYNLRRFMYCPPTLDGRAHGEHLHEQFRK
TILTSHAFFPYIKTRVNTHKETILTPIEVAIEDMOKTQELAFATHQDPADPKMLQ
MVLQSVGTTVNOGPLLEVAQVFSEIPSDPKLFRHHNKRLLCFKDFTKCEDALRNK
SLIGPDQKEYQRELERNYHRLKEALQPLINRKIPQLYKAVLPVTCHRDSFSRMSLRKM
DL"
BASE COUNT 781 a 458 c 509 g 704 t
ORIGIN
Query Match 34.6%; Score 2360.2; DB 9; Length 2452;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 2414; Conservative 0; Mismatches 28; Indels 7; Gaps 3;
QY 4367 ttgacaatatcaagagcagagattgaaacaggaagcattgatgtgaaacctggtacag 4426
|||||
Db 1 TTGACAAATCAAGAGCAGAGATTGAACACGAAAGCACTGATTGATGGAACCTGGCTACAG 60
QY 4427 aagcaaacctaatcatattagatacatagagatgtgttcagaccgttctgtaacgg 4486
|||||
Db 61 AAGCAAACTAATCATTTAGATAACATTAGAGATGTGTTCAAGACCGTTCTGTAACGG 120
QY 4487 aatccaaagagagcattcttggtytgagtgctaaagtgtcactacacagcatgycctgta 4546
|||||
Db 121 AATCCAAAGAGAGCATTTCTGGTGAGTGCCTAAAGTGCCTACTACACAGCATGGCCTGTA 180
QY 4547 acccaagtgcagttatctacacaacactgtttgtctacacagagagccttggtttcaagt 4606
|||||
Db 181 ACCAAATGCAGTTTATCTTACACACACTGTTTTCCTACACAGAGAGCCTTGTTCAAGT 240
QY 4607 ttctgtaactctatttgaagaagagacagagcagtgctgatttatgcctcaggcttc 4666
|||||
Db 241 TTCTGGAACCTTTATTGGAAGAAGAGACAGAGCAGTGTGCTGATTATGCTCAGGCTTC 300
QY 4667 tccgacactgtagcagtagcatcggtacatacagtcacacccagtgctcccttacc 4726
|||||
Db 301 TCCGACACTGTAGCAGTAGCATCGGTACAAATACGGTCACACGCGCAGTCCCTTACC 360
QY 4727 tactaatgagcacaaccttgagattggyaataactttgccagggttaaatgcaggtac 4786
|||||

Db 361 TACTAATGAGGCCAAACCTTTGAGATTGGGAATACTTTGCCAGGGTTAAATGCAGCTAA 420

QY 4787 caatgtcaactatcctccttggtggtgcacatctcagaatttaatgaagaattcttaagac 4846

Db 421 CAATGTCACTATCCTCCTTGTTGGGCACATCTCAGAATTTTAATGAAGAATTTCTTAAGAC 480

QY 4847 gtctcctaaagactatattgacatatgtcgaagaagatctggaattgagggaaacaacat 4906

Db 481 GTTCTCTAAGACTATATTGACATATGCTGAAGAAGATCTGGAATTGAGGGAACAACAT 540

QY 4907 ttccctgacaggtccagagatctggtttccaatctccatagatcttctgatactgtga 4966

Db 541 TTCTGTGATCAGGTCAGAGATCTGGTTTCAATCTCCATATGATCTTCTTGATACGTGA 600

QY 4967 aaatgaaggaacaccagagagatcctgaaatgttgtattgatctaatgtacagaattgcca 5026

Db 601 AAATGAAGGAACACCAGAGAGATCCTGAATGTGATGTGATCTAATGTACAGAATTGCCA 660

QY 5027 aggtttaccagacctctccaga---gcgattgacctggttgagaacatggtgcaggaagc 5083

Db 661 AGGGTTACCAGACCTCTCCAGATCTGCCATTGACCTGGTTGCAGAACATGGCCAGCAAGC 720

QY 5084 actcagaacggaagcaatcatctgtgaagctgcacagtgctagtccactcagcagcaacttg 5143

Db 721 ACTCAGAACGGAACCAATCATGTGAAGCTGCACAGTGTCTAGTCCACTCAGCAGCACTTG 780

QY 5144 ttgctgaatatltgagcatgtctggaagccggaatatcttcctgtggtggtgtgaacat 5203

Db 781 TTGTGATATTTGAGCATGTGAGAGACCAGAAATATCTTCTGTGGATGTGTAACAT 840

QY 5204 ttcaagaatatctcataatgttttagaagaatctgcggtctcagatgatgtgtatctc 5263

Db 841 TTCAGAAATTTTCATCTAATGTTTGAAGAATCTCGGCTCTCAGATGATGTGTAATCTC 900

QY 5264 cagatgaagaaggtatctgtctctggaanaatacttactagtcaggaactgttggtgttac 5323

Db 901 CAGATGAAGAAGGTATCTGCTCTGGAATAATACTTTACTGTGAGTCAGGACTGTGGGATTAC 960

QY 5324 tggaaacaagcagctgtctccttctctatggtgctgcatgtatgaagcagtttaagtgtt 5383

Db 961 TGGAAACAAGCAGCTGCTTCCCTTCTATGGCTGGCATGTATGAAGCAGTTAATGAAGTTT 1020

QY 5384 acaaaagtacttattcctatctcagtaagctaatacgggagtcaaaagaactatccacaattc 5443

Db 1021 ACAAAAGTACTTATTCCTATTCATGAAGCTAATCGGATGCAAAAGAACTATCCACAATT 1080

QY 5444 atgtaaacctcaagaagcatcagcaaaatgttcatcagagtaactggtggtgagcgga 5503

Db 1081 ATGTTAACTTCAAGAAGCATTCAGCAAAATTTGTCATCAGAGTACTGGCTGGAGCGGA 1140

QY 5504 tgtttggacctatttcgtgtgtgtttttatgaaaccaagttcggtgatttgatgaac 5563

Db 1141 TGTTTGGCACTTATTTCTGTGTGGTTTATGGAACCAAGTTCGGGGATTTGATGAAC 1200

QY 5564 aagaattgttttaacaaggacctgcaataaaccaaactgcagagatatctcacagaattgg 5623

Db 1201 AAAAATTTGTTTACAAGGAGCCTGCAATTAACCAAACTTGCAGAGATATCTCACAGATT 1260

QY 5624 agggatlttaacggagaagaatttggagaggaatgtgttgaagtaatacaaaacttaac 5683

Db 1261 AGGGATTTTACGGAGAAAGATTGGAGAGGATGTGGTGAAGTAATCAAAAGACTCTAAATC 1320

QY 5684 ctgtagacaagtgtaaattagatcctaacaaggcatatatcagaattacctaattgagac 5743

Db 1321 CTGTAGACAAGTGTAAATTAGATCCTAAACAAGGCATATATTCAGATTACCTATGTGAGC 1380

QY 5744 catacttgacacataltgagatgaaggaagaaatcaccaatttcgacaaaattacaatc 5803

Db 1381 CATACTTTGACACATATGAGATGAAGACAGAAATCACCTATTTCGACAAAATTTACAATC 1440

QY 5804 ttctgtgattcatgtactgtacaccccttactttagatggtcggtgccatggtggaacttc 5863

Db 1441 TTCTGTGATTCATGTACTGTACACCCCTTACTTTAGATGGCCGTGCCCATGGGGAACCTTC 1500

QY 5864 atgacaattcaaaaaggaaagaccattctctgactaagtcctcatgctcttccttataataaa 5923

Db 1501 ATGAACAATTCAAAAGGAAGACCATTTCGACTACGCTCATATGCCCTTTCTTATATTA 1560

QY 5924 caaggtcaatgtcaactcataaagaagaatcatcttaacaccaattgaagttgctatg 5983

Db 1561 CAAGGTCAATGTCACTCATTAAGAAGAGATCATCTTAACCACAATTGAAGTTGCTATTG 1620

QY 5984 aggcacatgcagaaaaaagacacagagatgtgcatltgcaacacatacagatcccgagacc 6043

Db 1621 AGGACATGCAGAAAAAGACACAGAGAGTTGGCATTTGCCAACACATCAGGATCCCGACAGC 1680

QY 6044 ccaaatgtctcagatggtlactccaggaatctgtaggcaccaagtgaatcaggggctt 6103

Db 1681 CCAAAATGCTTCAGATGCTACTCCAGGATCTGTAGGCACACAGTGAATCAGGGGCTT 1740

QY 6104 tggaaagtgcaccaggttttctgtctgaaataacctagtgaccocaaagctcttcagacatc 6163

Db 1741 TGGAAGTGCCCAGGTTTTTCTGTCTGAATAACCTAGTGACCCCAAGCTCTCAGACATC 1800

QY 6164 ataataaactgcgactctgctttaaagaatttactaaagaatgttgaagatgccttaaga 6223

Db 1801 ATAATAAAGTGCACACTGCTTTAAAGATTTTACTAAAAGGTGTGAAGATGCCTTAAGA 1860

QY 6224 aaaaatgaagcttaattgggctggttcaaaaggaglatcaaaagggaattg--gggaact 6281

Db 1861 AAAATGAAGCTTAATTGGCGCCGATCAAAAGAGTATCAAAAGGAACTGGAGAGAACCT 1920

QY 6282 atcttcgcttaaaaggagccctacagagatctcctcagtgcgaatgagcttcgcaaaa 6339

Db 1921 ATCATCGCCTTAAGAAGGCCCTACAGCCACTGATCAACAGAAAGATCCCTCAGTTATACA 1980

QY 6340 agccagtatgtctgtcccccgtccacagagatctcctcagtgcgaatgagcttcgcaaaa 6399

Db 1981 AGGCAGTATTCCTGTACCTGCACAGAGATTCCCTCAGTCGAATGAGCCTTCGCAAAA 2040

QY 6400 tggatctctaaactgaatgcaactgttttattcatctcgtcaaaagagccatglatccaacat 6459

Db 2041 TGGATCTCTAACTGAATGCACCTGTTTATTCATCTGCAAAAGAGCCATGTATCAACAT 2100

QY 6460 cgaagtgtgaaaagatctatttggaaaccacaatggaatggaatcttggaaattatlatca 6519

Db 2101 CGAGTGTGAAGAATCTATTGGAAAAACAGCATGGAATGGAATTCGGAATTTATYTCA 2160

QY 6520 ttgaagaatgcagtggtggccaagaatatcaaatgtlagatgtttaaagctltgagaatcag 6579

Db 2161 TTGAAGAATGCAGTGGCCAGAAATAATCAAAATGTAGATTGTTAAACGCTTGAGAATCATG 2220

QY 6580 gctatggttctaatgtltcgggtaaacaagctgtatctttaaagacattttaatgactca 6639

Db 2221 GCTATGTTCTTATATGTTCTGTAAACAAGCTGTATCTTTAAAGACATTTTAATGACTCA 2280

QY 6640 aaggtacactatatacatltaaccatttataccatagctlaaggttaaaaaatttattcact 6699

Db 2281 AAGGTACACTATATCATTTACCATTTATATACCATAGCTAAGGTTAAAAATTTATTCACT 2340

QY 6700 ttaagttcglattttttaattataacacatttataagattcatttggaccattttaa 6759

Db 2341 TTAAGTTCGTATTTTAAATTATATATACCATTTATAGATTCTTTGGAACCATTTTAA 2400

QY 6760 atgtagtaatgtcttattttaaaggtactaaaaaataatgtgaatgtttac 6808

Db 2401 ATGTAGTAATGCTTATTTTAAAGGTAATCTTAATAATATGTGAATGTTTAC 2449

RESULT 9

AX173175

LOCUS AX173175 6372 bp DNA linear PAT 03-JUL-2001

DEFINITION Sequence 1 from Patent WO0142295.

ACCESSION AX173175

VERSION AX173175.1 GI:14598036

KEYWORDS

SOURCE	human .
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	Lu, P., Garman, J.D. and Candia, A.F.
TITLE	Clasp-7 transmembrane protein
JOURNAL	Patent: WO 0142295-A 1 14-JUN-2001; Arbor Vita Corporation (US)
FEATURES	Location/Qualifiers
source	1. 6372
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/note="full length human CLASP-7 cDNA"
CDS	13. 6156
	/note="human CLASP-7"
	/codon_start=1
	/protein_id="CAC43731.1"
	/db_xref="GI:14598037"
	/translation="MAASERAFAHKINRTVAAEVKKOYRSRSGSPHSSRCSSSLG VPLTEVEPLDFEDVLISRPDAEPGLRDLVEFPADDELLLPRECRTEBGI PKD EKLDQVRAAVEMYIEDWYVHRRYQYLSAAYSPVTTDQREKQGLPRQYEQDASG DERSGPEDNSRSGSGSPEDTVPRSSGASSIFDLRNLAADSLPLSILERADEDVDR NETLRQHRPPALLTLYPAPDEDAVERCSRPEPPREHFQRI LKCLSLFEIETEP IFGI LALYDVREKKKISENFYEDLNSDSMKGLRAHETHPAISTLARSAIFSYTYPSP DIFLVIKLEKVLQOGDISECEPVMVLKEVDTA NKKEKLEKRLAAEQFCSTGlyRyM PFAATVAVHLANIVSSAGQIDRDSDEGERRPATMDRRRGPODRASSGGDACSFSGR PATLVTVNFEKQAERLSDEDLFKFLADMRPSSLRLRPVTAOLKIDISAPENPH ECLSPELLHIKYPDPGRGRPTKEILEFPAREVYAPHTSYRNLLVYPHSLNFSRQGS VRNLAVRVQYMTGEDPSQALPIYFGKSSCEFTREATPVVYHNKSPEFYEEKLILP ACVTENHHLLETFYHVSQCDRPGTALET PVGFTWIRLQHGRLRTGPFCLPVSVDQP PSYSLVLPDVALPGMRWVDFKGEFVELTAVSSVHPQDPYLDKFTLVHVEGAEF FRKLDTVLSEGNVEQELRASIALRLASPEPLVAFSHVLDKLVRLVIRPPIISQIV NLGCAFEMAHVSVLSVHRSLEAAQDARGHPQLAAVYVHAEFRLPGTEPSLBDGAPV TVQAATLARSGRPASLIYLA RSKSISSNPD LAVAPGSVDDEVSRILASKLHBEELAL QWYSSSAVREAILQHAWFFQLWVKSMAHLHLGRLDTPRKLRFPGRFLDITLALY GSVGLVITRVHKDVELAEHLNASLAEFLSDLLSLVDRGFVPSLVRAHYKQVATRLQS SPNPALLTLRMEFTRIICSHHHVVTNLPCCP LSPASPSPSVSSTTSQSTFESSQA PDPKVTSMFELISGPERQHFPLAGLLTELALALEPEAGFLLHKKAISAVHSLLCGH DTDPRYAEATVKARVAELIYPLLSIARDTLPRLHDFAEAGPGORSRLASMLDSTEGEG DIAGTNPASVAMAIAGGPLAPGSRASISOGPPTASRAGCALSAESSRTLACVYLWLK NTEPALIORWATDITLPOLGRLLDLYLCLAEFYKKKAFERINSLTFEKSIDMKAR LEEALIGTIGARQEMVRSRSRSPENPENVRWKSVTHWKOTSDRVDKTEMEHEA LVEGNLATEASLVLDITLEIIVQTVLMS EARESVLCAVLKVLVLSGSAQSLAFLOHG LATQALVSKPEELLFEEDTELCADLCRLLRHGSRI STIRTHASASLVLMRQNEE IGHNFA RVKMQV TMSLSLVGTTONFSEEH LRS LKTLITYAEEDMGLRDSYFAEQVO DLMFNLMILITDTVKMKHEHEDPEMLIDLMYRIARGOGSPDLRLTWLQNMAGKHAEL GNHAEAAQCVMHAALVAEYLALELDQRHLPGVCVSEFQNISSNVLEESAIISDILSPD EEGFCSGKHFTELGLVGLLEQAAAGFYTMGGLYEAVNEVYKNLIPILEAHYKKLAAV HGKLOEAFETKIMHOSSGWERVEGTYFRVGYGAHFGDLDQEFVYKEPSITFLAEISH RLEEFYTERFGDDVEI IKDSY PVDKSKLDSQKAIQIITYVEPYFDITELKDRVTFD RNYGLRTFELCTPFTPDGRAHGLPEQHKRKTLLSTDHAFYIKTRIRVCHREETVLT PVEVAIEDMOKKTRELAFATBODPPDAKMLQWVLOGSVGPTVNQGPLVAYQFLAEIP EDPKLEFRHNNKLRLCFKDFCKCEDALRKNKALIGPDQKEYHRELERNYCRREALQP LITQRLPQLMAPTPPGLRNSLNRA SFRKADL"
BASE COUNT	1273 a 2050 c 1849 g 1200 t
ORIGIN	
Query Match	31.2%; Score 2132.8; DB 6; Length 6372;
Best Local Similarity	60.4%; Pred. No. 0;
Matches 3863; Conservative	0; Mismatches 2262; Indels 273; Gaps 10;
QY	26 ccgagcgccgccttcgcccagaagatcagcagaacggtgcgagccggaagttagaagc 85
Db	23 CCGAGCGCGCGCTTCGCGCACAGATCAACAGGACGGTGGCCGCGCAGAGTGGCGAAGC 82
QY	86 agatctccgacataatagtggtctccccaactgctcaaaaaccttaatatgttgca 145
Db	83 AGGTGTCCCGGAACGAGTGCTCCCCCACTCCACGAGCGCTGCACGACGCTC----- 137
QY	146 atatacccatcacacacagtgcccttaccgaagcagtagatccagtgatttggaag 205

Db	138 -----CCTGGGGGTCCCACTGACTGAAGTTGTGAGGCCCTGGACTTTGAGG 184
QY	206 attaccttactcatcccttggctgtgtgattcttggcccttaccggatttgatgaat 265
Db	185 ATGTACTTCTGAGCCGGCCACCAGATGCTGAGGCCCGGCCCTCAGGGACTGTAGAAAT 244
QY	266 ttccctcagatgatataatgaagtgtttatagtccttcgggactgcagaactctgttcag 325
Db	245 TCCCACTGATGACTGTGAGCTGCTGCTGCAGCCCCGGGAATGCCGGACCGAGGCCG 304
QY	326 ctgtacctgaagaagltgaatggtatccacatgttagagactgtataagaagltatacag 385
Db	305 GGATCCCCCAAGGATGAAAACTGGATGCCCAGGTGAGGGCCCGGTGAGATGTATATTG 364
QY	386 aagactggccaattgtcatcagaanaatatcataaattggyaacagatlttaatccaata 445
Db	365 AGGACTGGGTCAATTGTCCACAGAAAGTATCAGTACTGAGTGACATACAGCCCCGTCA 424
QY	446 catlagataacagaagaagcaaaaggtttgccaanaacaagttttgaatctgatg 505
Db	425 CCACAGACACACAGCGGGAGCGACAGAAAGGGCTTCCCGCCAGGTCTTTGACGACGATG 484
QY	506 aagctccagatgycacacagctaccagatgatcaagatgaaccttaagaagcttcaatgt 565
Db	485 CTTCGTGAGAGACGAGAGGTCCGGCCCTGAGGACTCGAATGACTCCCGCGCTCGGGCT 544
QY	566 caatagatgataccccaagggtagctggccctgtagtatctttagcttgaaaaaattcac 625
Db	545 CCCCCGAAGACACCCCTCGAAGCAGTGGTGCCCTTAGCATCTTCGACCTGAGGAACCTGG 604
QY	626 ttccctgatgcttgccttcccaatttacttgatgcgaactccaatgaaagaaatagaacgctc 685
Db	605 CAGCTGACTCATTTGCTGCCCTCTCTGCTAGAGCGGGCGGCCCCAGAAAGATGTGACCGGC 664
QY	686 agaatgatgaccaaagaaatcaaacggtcacaaagaacttttgccttgcatccatcac 745
Db	665 GCAATGAAACCCTTGACGGCAGCACCGGGCCCGCCCTGCTCACCCCTGACCCGGGCAC 724
QY	746 cagatgaggaagaaccaatagaaacggttagtggttcctgatatacccaagaacatttg 805
Db	725 CTGACGAGGATGAAGCCGTGGAACGCTGTAGCCGCCAGACCAACCCCGGACACTTTG 784
QY	806 gtcaaagacttctgtiaaaatgcttatcactcaagtttgaatltgaatltgaaccattt 865
Db	785 GACAAGGATCTTGGTCAAGTGTCTGCGTCAAGTTCGAGATTGAATATGAGCCCATCT 844
QY	866 ttgcaagtttgcttataatgatgttcaaggaagaanaagaatttcaganaacttttatt 925
Db	845 TTGGGATCTTGGCTCTGTATGATGTGCGGGAGAAAAGAAGATCTCGGAGAACTTCTACT 904
QY	926 ttacaccttaattctgagcagatgaaaggtgtgtacgtccacatgttaccactgctgcca 985
Db	905 TCGACCTGAACTCGGACTCGCATGAAAGGGGCTGCTTGGGGCTCATGGCACCCACCTGGCCA 964
QY	986 ttactaccctgycgaagatcagcaaatltttctatcactatccttcccaagatgttttc 1045
Db	965 TCTCCACCCCTGGCCCGCTCTGCCATCTTCTGTGACCTTACCCCTCACCTGACATCTTCC 1024
QY	1046 ttgtaataaagctagaaaagctctacagcaagagacatltgagagtggtgcagaacct 1105
Db	1025 TGTTCATCAAGTTGAGAGAAGTGTCTTACGCAAGGGGACATCAGTGAGTGTGTGAGCCTT 1084
QY	1106 atatgatlttcaagaagcagatgcccaccaagaataaagaaaacttggagaaaactgaaga 1165
Db	1085 ACATGCTGTGAAAGAGTGGACACACAGCCACAGAACAAAGAGAGCTAGAGAACTGCGCC 1144
QY	1166 gtcaagcagatcagtttgcccaagacttgggaaataatcgatgccttctgttgactg 1225
Db	1145 TGGCGGCCGAGCAGATTCTGCACCCCGCTGGCGGCTTACCGCATGCCCCCTTGAGCGG 1204
QY	1226 caatccalltaatgaatatgttagcagtgcttgaggaglttggaaagagattctacagaag 1285
Db	1205 CCGTGCACCTTGGCCAACATCGTGAAGCAGCGCTGGGCAAGCTGGAACCGGACTCTGACTCGG 1264

QY	1286	tagaatacagtactctggaagaacgaaaaagygctcttggtcagagaggaaggaattctagtattg	1345
Db	1265	AG-----GCGGAGCGCGCGGCGACGCTTGACAGACCGCGCG-----	1298
QY	1346	ttgycagacgatacacttgaagaagacacaagaatgagatgatgtctgttaacttgcagagct	1405
Db	1299	-CCGTCGGGGCCCCAGGACCGGGCGAGTAGTGGGGACGACGCCCTGCAGCCTTCTCTGGCT	1357
QY	1406	ttcgaccagctactctcacagtgacaaattttttaagcaggaagagacccgttaagtgtg	1465
Db	1358	TCCGTCACGCCACCGCTTAACGTGCACAAACTTCTTTAAGCAGGAGGCGTGAAGCGACTCAGTG	1417
QY	1466	atgaagatctctacaaattcccttgctgatatggaagggccatctctgtcttacgycgac	1525
Db	1418	ACGAGGACCTCTTCAAGTCTCTGGCTGACATGAGGCGCGCGTGTCCCTGTGCGGCGAC	1477
QY	1526	taagacctattacagctcagctcagaatagacatttctccgcacctgaaatccccatt	1585
Db	1478	TACGTCTGTGACTGCCCCAGCTCAAGATCGACATTTCTCCGGCTCCTGAANAATCCCCACT	1537
QY	1586	attgcctaactccggaagctgtctcaagtgaagctttacccctgacagtagagttlagaccta	1645
Db	1538	TCTGCCCTCTCCCTGAGCTGCTTCATATCAAGCCCTAACCCGGACCCCAAGGGCGCGCCA	1597
QY	1646	ccagagaaatcttagagtttccgcgaaggagatgttatgttccaacaactactacagaa	1705
Db	1598	CCAAGAGATTCTGAGTTCGCCCGCGCGGAAGTCTATGCCCCCATACCACTACAGTAACAGGA	1657
QY	1706	atcttctctacataaacctcagagctctaattttgccaatcgttcaaggtctcgtagaa	1765
Db	1658	ACCTGTCTAGCTGTATCCCGCACAGCCTCAACTTCACAGACCGCGGCTCCGTGCGCA	1717
QY	1766	ataaacagtgaaagtcagtttatgtatgagagagatccaagcaatgccaatgcccgttaa	1825
Db	1718	ACCTGTCTGTGCGAGTGCAGTACATGACAGCGGAGAGACCCACAGCGCTCTGCCGTCA	1777
QY	1826	tctttgttaaatcttagctgtttcagaattttccaaggaagccctatacagccgtagtatac	1885
Db	1778	TCTTTGGCAGAAGTCCAGCTGCAGTGAATTTAACCCGGAGAGGCTTCACACCGGTGTCTACC	1837
QY	1886	ataacaggtctcctgaattttcatgagaataccaaggttaagcttctcgtctactttaactg	1945
Db	1838	ATAACAAGTCCCCCGAGTTCTACGAGAGTTCAAGCTGCATCTTCCAGCCTGCGTGACAG	1897
QY	1946	accatcatcacttgcttttttacttttatacatgtaagttgttcaacaaaacaatactc	2005
Db	1898	AGAACCATCACCTGCTGTTCACCTTCTACCATGTCAAGTGCACGCCCGCGGGGCACTG	1957
QY	2006	ctcttgaacacaccagttggtatatacatgtataccaaatgcttcagaatggaaggttgaaga	2065
Db	1958	CCCTGGAGACACCCGTGGGCTTTACTTGGAATCCCACTGCTGCAGACGGGCGCCTTAGGA	2017
QY	2066	ctgcccagtttgtcttgcagctcctcatgtggaaaaaccaaccaaggttatctgtactgt	2125
Db	2018	CCGGCCCTTCTGTCTCCAGTGTCTGTGGACCAAGCCGCGCCAGCTATTCCGTGCTCA	2077
QY	2126	ctctcaggttctctactaccttgcacatgaaatggttagataatacacaaggtgttttaatg	2185
Db	2078	CACCCGATGTGCGGCTTCCGGGCATGCGCTGGTGGACGGTCAACAAGGCGTGTTCAGTG	2137
QY	2186	ttgaagttgttgtgtttcgtctatccatacacagaatccctatcttgacaaatttttg	2245
Db	2138	TGGAAGCTACACAGCCGTTCTCTGTGCACCCCAAGACCCCTTACCTGGACAAATCTTCA	2197
QY	2246	ctctgttcaatgtctcttgaatgaacacactgtttccagtcggaattggygacatgcyaatca	2305
Db	2198	CCCTGTGTCACAGTCTCTGGAGGAGGAGGAGCCCTTCCCATTTCCGGCTCAAGGACACTGTGTGA	2257
QY	2306	tggaataaacttagaaaaatgaattgaagagcagtatattcagcactgaattcatcccaagc	2365
Db	2258	GCGAGGCAACGTGAGCAGGAGCTGCGGGCCAGTCTTTCAGCACACTGCGCCTGGCCAGCC	2317

QY	2366	tggaaccagtggtccgatttcttcaatctctctgtagataaaactgatactttagttatta	2425
Db	2318	CCGAACCCCTGTGGCCTTCTCCACCACAGCTGCTGGACAAGCTCGTGCGTCTGGTCATCA	2377
QY	2426	gaacctctgcatctgctggccaatatagtltaacctagggtcaagcatctlttgaagccatg	2485
Db	2378	GGCCCCCGATCATCAGTGGCCAGATTGTGAACCTGGGCCGTGGAGCCTTTGAAGCAATGG	2437
QY	2486	catcaattataaatcgacttcacaaaacttgyaagaaatcatgaccagcatgycagaa	2545
Db	2438	CCCATGTAGTCAGCCCTTGTTCACCCGGAGCCTGGAGGACCCACAGATGCCCGGGTCACT	2497
QY	2546	acagccttcttgcatcatatatattcatgttttccgcctaccaataacttaccctaatt	2605
Db	2498	GCCCCACAGCTGGCTGCCCTACGTCCACTACGACCTTTCGCCCTTCTGGCACTGAGCCCA	2554
QY	2606	catcatcaccaggtccctgggggttgyagagatcagtgcatatgcccacaatggtagat	2665
Db	2555	-----GCCTCCCGGATGGGGCCCCCTCCAGTAGACAGTGCAGGCTGCCACACTGGCCCTG	2608
QY	2666	ctgcgylgagacctgcaagccttaattaaatcgltctcgaaagcctlagtaatagaatc	2725
Db	2609	GCTTGTGTGCCCCCGCAAGCCTCTACCTGGCGGTTTCCAAGAGCATCAGCAGCAGAAC	2668
QY	2726	cagatatatctggyactcccaacgctcaccagatgatgaagttcgatcaatcatcggyagta	2785
Db	2669	CTGACCTCGCCGTGGCCCCCTGGCTCTGTGGATGACGAGGTTTCCCGCATCTGGCCAG--	2726
QY	2786	aggytttagatcgctccaattcctctggtttaacactgtygtlccaaaagctgcccacatggy	2845
Db	2727	-----	2726
QY	2846	gatccaaccacgtccaagtgcagaatcaacacagctatgatacgaaattgtaatcgta	2905
Db	2727	-----	2726
QY	2906	tgctctcgcacacagagacgltcaagtlcttacaacaaltaacgygaagctlaaccaacta	2965
Db	2727	-----	2726
QY	2966	aaaagcttttcacgagagagctggtcttgcagtggtgtttgcagtgycagcgttcggg	3025
Db	2727	-CAAGCTGCTTCACGAGGAGCTGGCTCTGCAGTGGGTGTGACGACGACAGTGCCTACGCG	2785
QY	3026	aatcagctttgcacaagaagcctggttcttttttgtaattatggttaaagagcatggtgcacc	3085
Db	2786	AGGCCATCCTCCACGACGCGCTGTCTTCTTCCAGCTCATGTGTAAGATATGGCGCTGC	2845
QY	3086	attataactttaatgataaacttgaggctccaaagaaaagtcgttttccagaacgtttca	3145
Db	2846	ACCTGCTGCTTGGCCACGCGACTAGACACACACCCCGCAAGCTGGCGCTTCCCGGACGCTTCC	2905
QY	3146	tggatgacattgcaagctctgttcagcaacgattgctagtgatatagtltcaagattcaga	3205
Db	2906	TGGACGACATCACTGCGCTTGGTGGGCTGTGTGGCCCTGAGGTGCATCACCCGTGTCCACA	2965
QY	3206	aggacacagaaatggttggagagactcaataacaagccttgcattcttctcaatgatctgt	3265
Db	2966	AGGATGTGGAGCTGGCCGAGACCTCAACGCCAGCCTGGCTTCTTCTCAGTGACCTTC	3025
QY	3266	tgctcgttatgacagagagattgttttagccttaataaagtcctgctataaacaggtgt	3325
Db	3026	TGTCCCTGTGTGACCGGGGCTTGTCTTCAGCCTGTGTCCGGGGCCCACTACAAGCAGGTGG	3085
QY	3326	cttcaaaagccttactcaatlaaccgaatccocagtgttctggtgtccttgtagcgtgatctc	3385
Db	3086	CCACGGGGCTCCAGTCTGTCCCTTAATCCAGCAGCCCTGTCTGACCTTGGCATGGAAITCA	3145
QY	3386	tacgaatcatctgcagtgcatgagcaactatggttacatlaaacttaaccctgcagcttaacta	3445
Db	3146	CCCGCATCTGTGTGACGCGCACGACACTACGTGACCCCTCAACCTCCCTGTGCCCCCTGT	3205
QY	3446	ctccacctgcatactccatcaacctctgttcttctctgcaacatactcaagattctgatat	3505

Db 3206 CACCTCCAGCCTCGCCCTCCCTCTGTGTCTCCACCACCTCCAGAGCTCCACCTTCT 3265
QY 3506 ctacgaaatgtacaagaccaaaagattgcaaatatttgaaatataccgtgccttcgcgc 3565
Db 3266 CCAGCCCAAGCCCCGGACCCCAAGGTGACCAGCATGTTCCAACTGAGTGGAACATTCGGGC 3325
QY 3566 aacagcattatttggcagacctgtgttaacagagctggctgtcaattttagacctgatg 3625
Db 3326 AGCAGCATTCTCTAGTGGCTCCTGTGACGGAGCTGGCACTGGCCCTCGAACCTGAGG 3385
QY 3626 ctgaaggactgtttgatgtcataaagaatcatcaatatgtgtacacaatttactctcca 3685
Db 3386 CTGAAGGGGCATTCTGTGTGCACAAGAAGGCCATGCTGTGTGCACAGCCTGCTATGTG 3445
QY 3686 gtcaagcactcagaccgcggtactctgaccctcagataaagctcgaagtgcagtgccatgtt 3745
Db 3446 GCCATGACACTGACCCCCGCTACGCCGAGGCCACTGTGAAGGCTCGTGTGGCGAGCTGT 3505
QY 3746 atctacctctgattgtgtattatcatgtgaaacctgtacacctcagctgtatgtattacagaaa 3805
Db 3506 ACCTGCCACTGCTTTCGATTGCCACGGGATACCTTGCCACGGCTGCATGACTTTGCTGAGG 3565
QY 3806 ctcaacaatcaacgaggaagac-----caattgtatagccaactgatatgtaaaagt 3859
Db 3566 GCCCAGGTACGCGGTCAAGACTGGCCTCAATGCTTGACTCAGACACAGAGCGGAAGGGG 3625
QY 3860 aagacggaagtatgataagccagaccgttgcacatgcgaatcgcagggacatcgctccctc 3919
Db 3626 ACATTGGGGTACCATCAACCCCTCTGTGGCCATGGCCATTGCTGTGGGCCCTTAGCCCC 3685
QY 3920 aactaacaagcgctgcagttccctcctcactcagtcacacgagtgcgacaacactacct 3979
Db 3686 CTGGCTCCCGGGCAGCACTTCCAGGGGCCACCAACGGCTTCTCGCGCAGGCTGTGCC 3745
QY 3980 ttccagcagaatcaagtcgaagccttttgatctgtctacttttgggttctcaaaaatgcag 4039
Db 3746 TCTCTGCTGAGTCAAGCCCGACCTTGCTGGCGGTGTGTGCTGTGGTGCTGAATAAACACCG 3805
QY 4040 atgaaacagttctacagaagtgtttacagatctctcagtccttcgacgttaaacccggtat 4099
Db 3806 AGCCGGCGCTCCTGCAGCGCTGGGCCACTGACCTGACACTCCCCCAGCTGGGACGCTCTGT 3865
QY 4100 tagatctgttatctctgtgtgtctgtcgtttagtataaagggaanaagtgtttgaac 4159
Db 3866 TGGACTTGTCTGTAACCTTGTGCTGTGCTGCTTGTGATCAAGGGGAAAAAGCCCTTGAAC 3925
QY 4160 gaatgaatagcttgacctttaagaatatcaaaagacatgagcaaaagcttgaagaacta 4219
Db 3926 GCATCAACAGCCTCACATTCAAAAATCTCTGATATGAAGCGCGCTAGAGGAAGCCA 3985
QY 4220 ttcttggagcatagtggtccagcgcaagaatgtracgycgaagccgagacagctcgaga 4279
Db 3986 TTCTGGGTACCATCGGAGCTCGACAAGAATGTTCCGGCGAAGTCGTG----- 4033
QY 4280 gaagcccatctggaagtgcctttggaagtcaagaanaatttgaggtggaagaaagatatga 4339
Db 4034 -----AGAGGAGCCCGTTTGGGAATCCAGAGAACGTGCGCTGGCGGAAGCGCTCA 4084
QY 4340 ctcaactgycgtcaaaacagagaagccttgacaaaatcaagagcagagatgaaacaggaag 4399
Db 4085 CACACTGGAAGCAAACTCAGACCGCGGTGACAAAGACCAAGGATGAATGGAACACGAGG 4144
QY 4400 caactgatgtgaaacctggtctacagaagcaaacctaatacattttagatacattagaga 4459
Db 4145 CTTTGGTGAAGGAACCTTGCAACCGAGGCAAGCCTTAGTGTCTTGACACACTGGAGA 4204
QY 4460 ttgttgttcagaccgtttctgttaacgnaatccaagaagagcatcttctgtgagtgctaa 4519
Db 4205 TCATCGTGCAGACGGTGAATGCTTTCAGAAAGCCCGGAGAGAGCGTCTTGGGGCAGTGCTGA 4264
QY 4520 aagtgtactacacagcatgycctgttaaccaagtgcagtttatctacaacactgttttg 4579

Db 4265 AGTTGTGCTGTACAGCCTTGCGCAGTGGCCAGAGTGGCCCTCTTCTTGACAGCATGGCCTGG 4324
QY 4580 ctacacagagagcccttgytttcaaaagtttccctgaacctcttatttgaagaagagacagagc 4639
Db 4325 CCACCCAGAGGGCCCTTGTGTCCAAGTTCGCCGAGAGCTGCTGTTCGAGGAGGACACGGAGC 4384
QY 4640 agtgtctgatttatgcctcagcttctccgacactgttagcagtagacatcgtataacatc 4699
Db 4385 TGTGTGCCGACCTGTGCTTGAGGCTCCTACGACACTGTGTGGCAGCCGATCAGCACCATCC 4444
QY 4700 ggtcacacccccagtgccctcccttaacctactaaatgagcaaaaactttgagatggyaata 4759
Db 4445 GCACGCACGCCAGCGCCTCGCTGTACTCTGCTCATGCGACGAACCTTGAGATGGCCACA 4504
QY 4760 accttgcagggttaaaatgcaggtaccaatgctcactatccctctgtgtggyacatctc 4819
Db 4505 ACTTTGCCCGTGTGAAGATGCAGGTCAACCATGTCTCTGCTCCCTGTGGTGGGACGACGC 4564
QY 4820 agaatttaatgaagaattcttaagcgttctctaagactataatgacatatgtcgaag 4879
Db 4565 AGAACTTCAGTGAAGAGACACCTGCGACGTTCACTCAAAACCATCCCTACCTATGCTGAGG 4624
QY 4880 aagatctggaatttgaggyaaacaacatttccctgatacaggtccagatcgtgtttcaatc 4939
Db 4625 AGGACATGGGGCTTGGCGGACAGCACCTTCGCAGAGCAGGTCCAGGACCTGATGTCAACC 4684
QY 4940 tccatatgatcttctctgatactgtgaaaatgaaagaaaccaggaagatccctgaaatgt 4999
Db 4685 TGCACATGATCTCTGACGGACACGGTGAAGATGAAGGAACCAAGAGACCTTGAGATGC 4744
QY 5000 tgaatgataatgttacagaattgccaaggttacagaacctctcaga--gcgaattga 5056
Db 4745 TCATGCACTCATGTACAGAATGCCCCGGGGCTACAGGGCTCACCGGACTTTCGGCTGA 4804
QY 5057 cctgtgtcagaacatgycgcgaagcaacactcagaaagcaagcaatcatgctggaagctgac 5116
Db 4805 CCTGGTTGCAGAACATGGCCGGGAAGCACGCGAGCTGGCAACACGCGGAGCCGCC 4864
QY 5117 agtgtcagttccactcagcagcactgttgtcgtgaatatttgagcatgctggaagccgya 5176
Db 4865 AGTGCAATGTGCACGCGCGCCGCTCGTGAGTACCTCGCCCTGCTCGAGGACCAGC 4924
QY 5177 aatatcttccctgttggatgtgttaacatttcagaaatatctcatatgttttgaagaat 5236
Db 4925 GCCACCTGCCCCGTGGGCTGCCGTTTCTTCCAGAAACATCTCATCAACGTGCTAGAGGAGT 4984
QY 5237 ctgcggtctcagatgatgtgtatctccagatgaagaagtlatcgtctcgtgaaaaact 5296
Db 4985 CCGGCATCTCCGACGACATCTGTGCCCCGACGAGGAGGGCTTCTGTCCGGGAAGCACT 5044
QY 5297 ttactgaagtcagacttgttggattacttggaaacagcagctgtcctctctatagctg 5356
Db 5045 TCACGTAGCTGGGGCTGTAGGGTTGCTGGAACAGCGACCGGCTACTTACACATGGGCG 5104
QY 5357 gcatgtatgaagcagttaatgaagtttacaagtaacttccatattcatatgaagctaadc 5416
Db 5105 GGCTTCACGAGGCGGTGAATGAGGTCTACAAAGAACCTCATCCCATCTCGAAGCCACCC 5164
QY 5417 gggatgcaaaagaactatccacaattcatatgtttaaacttcaagaagcatctcagaaaaat 5476
Db 5165 GTGACTACAAAGAAAGCTGGCCCGGTGCACGCGCAAACTGCAGGAGGCCTTACACCAAGATCA 5224
QY 5477 ttcatagaagtlactgcttgggaagcgyatgttttggcaactatttcgttgtttttatg 5536
Db 5225 TGCACAGAGTTCGGGTGGGAGCGCGGTGTTCGGGAGGTATTTCCGGGTGGCTTCTACG 5284
QY 5537 gaaccaagttcogggaatttggatgaacaagaatttgttttaacagagagccttcaataacca 5596
Db 5285 GCGCCCACTTCGGTGACCTGAGTGAAGCAGGAGTTTGTGTACAAAGGACCAATGCATCACGA 5344
QY 5597 aacttgcagagatatctcacagattggaagatttaccggaagaagatttggagaagatg 5656
Db 5345 AGCTGGCAGAGATCTCACACCGGCTTGAGGAGTGTCTACACGGAGAGATTTGGCGACGACG 5404

QY	5657	tggttgaagtaatcaaaagactctaattccctgtagacaagtgtaaattagatccctaacaagg	5716
Db	5405		5464
		TCGTTGAGATTATCAAGACTCTTACCCTGTGGACAAGTCCAAGCTTGACTCACAAAAGG	
QY	5717	catatattcagattacactatgtgtagccatactttgacacatatgagatgaagacagaa	5776
Db	5465		5524
		CCTACATCCAGATCAAGTATGTGGAACCGTACTTTGATACCTACGAGCTCAAGACCGGG	
QY	5777	tcacctatttcgacaaaaattacaatcttcgtcgtatcattgactgtacaccccttact	5836
Db	5525		5584
		TGACCTACTTTGACCCGCAACTATGGGCTTCGCACATTCCTGTTCTGCACGCCGTTCA	
QY	5837	tagatggccgtgcccattggggaacttcataaacaattcaaaagaagacattctgacta	5896
Db	5585		5644
		CGGATGGCGCGCACAGGGGAGCTGCCCGACACACAAGCGTTAAGACGCTGCTCAGCA	
QY	5897	cytctcatgccttccttatataataaacaagggtcaatgtcactcataagaagatca	5956
Db	5645		5704
		CCGACCAAGCCTTCCCTACATCAAGACTCGCATCCGTGTGTGCCACCGGAGAGACGG	
QY	5957	tcttaaccacaattgaagtgtctattgagacatgcagaaaaagaacacagaggtgcat	6016
Db	5705		5764
		TGCTGACGCCAGTGAAGTGGCCATCGAGGACATGCAGAAGAAGACACGGGAGCTGG	
QY	6017	ttgcaacacatcagatcccgccagaccccccaaatgcttcagatggtactccaaggatctg	6076
Db	5765		5824
		TTGCCACCGAGCAGGACCCACAGATGCTAAGATGCTACAGATGGTGTTCAGGGGCTCTG	
QY	6077	taaggaccacagtgaaatcaaggggccttggaaagtggcccagggttttctgtctgaatac	6136
Db	5825		5884
		TAGGGCCCAACCGTGAACCAAGGGTCCCTGGAGGTGGCCAGGTGTTTACAGAGATCC	
QY	6137	ctagtgacccaaaagctcttcagacataataaactgcgaactcgtctttaagaattta	6196
Db	5885		5944
		CGGAAGACCCCAAGCTCTTCCGGCATCACAACAAAATTGCGGCTCTGCTCAAGGACTTCT	
QY	6197	ctaaaaagtgtagaagatgccttaagaaaaataaagagcttaattggcgcggttcaaaag	6256
Db	5945		6004
		GCAAGAAATGTGAGGATGCGCTGCGGAAAAAATAAGGCCCTGATTGGCGCGACCAAGAAG	
QY	6257	agatcaaaagggaatg--gggaaactatcttcgccttaaaagagcgccctacagccctaga	6314
Db	6005		6064
		AGTACCACCGGTGAGCTGGAGCGGCAACTACTGCCGCCCTGCCGGAGGCTCTGCAGCCCTGC	
QY	6315	tcacagaaagtcctcctcattatccaagccagatgtgtccctgcacacagagattcc	6374
Db	6065		6120
		TTACCCAGCGGCTGCCCCAGCTGATGGCACCCACCCC---ACCCGGCTTCAGGAACCTCC	
QY	6375	ttcagtcgaatgagcttcgcgcaaaatggtatcttaaac	6412
Db	6121		6158
		TTGAACAGAGCAAGTTTCCGAAAGGCAGACCTGTGAGC	
RESULT	10		
AK055401			
LOCUS	AK055401	2607 bp	mrna
DEFINITION	Homo sapiens cDNA FLJ30839 fis, clone FEBRA2002429.		linear
ACCESSION	AK055401		PRI 31-OCT-2001
VERSION	AK055401.1	GI:16550120	
KEYWORDS	oligo capping; fis (full insert sequence).		
SOURCE	Homo sapiens fetus brain cDNA to mRNA, clone_11b.FEBRA2		
	clone:FEBRA2002429.		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.		
REFERENCE			
AUTHORS	1 (sites)		
	Kawakami,B., Sugiyama,A., Takemoto,M., Sugiyama,T., Irie,R.,		
	Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y.,		
	Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H.,		
	Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K.,		
	Wagatsuma,M., Murakawa,K., Kanehori,K., Takahashi-Fujii,A.,		
	Oshima,A., Suzuki,Y., Sugano,S., Nagahari,K., Masuno,Y., Nagai,K.		

		TITLE	and Isogai,T.
		JOURNAL	NEDO human cDNA sequencing project
		REFERENCE	Unpublished
		AUTHORS	2 (bases 1 to 2607)
		TITLE	Isogai,T., Otsuki,T. and Sugiyama,T.
		JOURNAL	Direct Submission
		COMMENT	Submitted (24-OCT-2001) Takao Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:genomics@hri.co.jp, Tel:81-438-52-3951, Fax:81-438-52-3952) NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: RAB and HRI.
		FEATURES	
		source	Location/Qualifiers
			1..2607
			/organism="Homo sapiens"
			/db_xref="taxon:9606"
			/clone="FEBRA2002429"
			/tissue_type="brain"
			/clone_lib="FEBRA2"
			/dev_stage="fetus"
			/note="cloning vector: pME18SFL3"
			35..1933
			/note="unnamed protein product"
			/codon_start=1
			/protein_id="BAB70917.1"
			/db_xref="GI:16550121"
			/translation="MAERRAFAOKISRTVAAEVRKQISGYSGPQLKLNINIGNIS
			HTTVPLEAVDPVDLEDYLITPHLAVDSGLRDLEFPDLEIYVSPRDCRTLVSA
			VPEEEMDPHVRCIRSYTEDEWAIVRKYHKLGTGFNPNTLDKOKERQGLPKQVFEES
			DEAPDNGSYDDQDDDLKRMSIDTTPRGSWACISFDLKNSLDALLPNLDRTPNEE
			IDRONDQKRSNRHKELEFALHPSPDEEPIERLSVPDIPKEHGQRLVLKLSLKEI
			EIEPIFASLALDYVEKKKISENFYFDLNSQMKGLRPHVPAAITTLARSAIFST
			YPSODVFLVIRLEKVLQOGDIGECAEPMIFKEADATKNKEKLEKLSQADOFQORLG
			KYRMPFAWTAIHLMNIVSSAGSLERDSTEVEISTGERKGSWERRNSSIVGRSLERT
			TSGDACNLTSFRPATLTVTNFEKQEGDRLSDLEDLYKFLADMRRPSSVLRRLRPTAQ
			LKIDISPAPEPNHYCLTPELLQVKLYPDSRVRPTRILEFPADVVPNTTYRNLLYI
			YPSLNFANRQGSARNITVVKQFMYGEDPSNAMSPSCTHPPIVVRPLPLWCYSIFS
			LLSWDFIT"
		BASE COUNT	843 a 500 c 506 g 758 t
		ORIGIN	
		Query Match	26.7%; Score 1821; DB 9; Length 2607;
		Best Local Similarity	100.0%; Pred. No. 0;
		Matches 1821; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	1	gtcgccgtcgccgacgacgcatggccgagcgccgcttcgcccagaagaatcagcaga	60
Db	14	GTGCGCGTGGCGCGACGACCATGGCGCGCGCGCTTCGCCAGAAAGATCAGCAGA	73
QY	61	acggtgcaagccgaagtaggaagcagatctccggaacaatatgtgtctccccaactg	120
Db	74	ACGGTGGCAGCCGAAGTTAGGAAGCAGATCTCCGGACAATATAGTGTCTCCCAACTG	133
QY	121	ctcaaaaaccttaatatgttggcacaatatatcccatcacaccacagtgccctaccgaa	180
Db	134	CTCAAAAACCTTAAATATGTGTGGCAATATATCCCATCACACACAGTGCCCCCTTACGAA	193
QY	181	gcagtagatccagtgatttggaagattacctcatctactcatccttggctgtgatctc	240
Db	194	GCAGTAGATCCAGTGGATTGGAGAATTACCTCATTTACTCATCTTTGGCTGTGGAATTCT	253
QY	241	gggccttaacgggattgattgaatttcctccagatgatattgaagttgtttatagtcct	300
Db	254	GGGCTTTACGGGATTTGATTGAATTTCCCTCCAGATGATATTGAAGTTGTTATAGTCCT	313
QY	301	cgggactgcagaactctgtttcagctgtacctgaagaagtgaaatggtatccacatgtt	360

Db	314	CGGGACTGCAGAACTCTTGTTCAGCTGTACCTGAAGAAAGTGAATGGATCCACATGTT	373
QY	361	agagactgtataagaagtatacagaagaactgygcaattgtcatcagaataatatacataa	420
Db	374	AGAGACTGTATAAGAGTTATACAGAAGACTGGCAATTGTCATCAGAAAATATCATAAA	433
QY	421	ttgggaacagatttaatcccaatacatatagataaaccagaagaagycaaaaggtttg	480
Db	434	TTGGGAACAGGATTTAATCCCAATACATTAGATAAACAGAAAGCAAAAAGGTTTG	493
QY	481	ccaaaacaagtttttgaatcctgataagctccagatgycaaacagctaccagatgatcaa	540
Db	494	CCAAAACAAGTTTTGAATCTGATGAAGCTCCAGATGGCACACAGCTTACCAGATGATCAA	553
QY	541	gatgaccttaaaagacgttccaatgtccaatagatgatatacccccaaggggtagctgygcctgt	600
Db	554	GATGACCTTAAAGACGTTCAATGTCAATAGATGATATCCCCAAGGGGTAGCTGGGCTGT	613
QY	601	agtatccttgacttgaaaaattcacttcctcgtatgcttgccttcccaattactgatcga	660
Db	614	AGTATCTTGACTTGAAAAATTCACTTCCTGATGCTTGTTCCTCCCAATTTACTTGATCGA	673
QY	661	actccaatagaanaatagaccgttcagaatgatgaccaagaagaatcaaacctcacaaa	720
Db	674	ACTCCAATGAAAGAAATGACCGTCAAGATGATGACCAGAAAGAAATCAAAACGTCACAAA	733
QY	721	gaacttttgccttgcatacaccagatgaggaagaaccaatagaacgcttagtgtt	780
Db	734	GAACTTTTTGCTTGGCATCCATCACCAGATGAGGAAGAACCAATAGAACGGCTTAGTGTT	793
QY	781	cctgatatacccaagaacattttgtccaagaactcttgtaaaaatgcttatcactcaa	840
Db	794	CCTGATATACCCCAAGACATTTTGGTCAAAAGACTCTTGTAATAATGCTTATCACTCAAG	853
QY	841	tttgaaatlgaaatlgaacccatttttgcaagtttgccttataatgatgtcaagaaaa	900
Db	854	TTTGAATTTGAATTTGAACCCATTTTGGCAAGTTGGCTTATATGATGTCAAGGAAAAAG	913
QY	901	aaaaagaatlccagaaaaactttatatttgacctaatlctgagcagatgaaaggttgtta	960
Db	914	AAAAAGATTTGAGAAAACTTTTATTTTGACCTTAATCTGAGCAGATGAAAGGGTTGTTA	973
QY	961	cgtccacatgtlaccaccctgtgcatactactaccctgccaagatcagcaatttttctatc	1020
Db	974	CGTCCACATGTACCACCTGCTGCCATTACTACCCTGGCAGATCAGCAATTTTTCATC	1033
QY	1021	acttatccctcccaagaatgttttcttgttaataaagctagaaaaaagctctacagcaaga	1080
Db	1034	ACTTATCCTTCCCAAGATGTTTTCTTGTAAATAAGCTAGAAAAAGTCCCTACAGCAAGGA	1093
QY	1081	gacatgtgagagtggtgcagaaccatatatgatatttcaaaagacagatgcccaccaagaat	1140
Db	1094	GACATTTGGAGAGTGTGCAGAACCATATATGATTTTCAAGAAGCAGATGCCACCAAGAAAT	1153
QY	1141	aaagaaaaactgagaaanaactgaaagatgcaagcagatcaagttttgccaagaacttgygaaa	1200
Db	1154	AAAGAAAAACTGGAGAAACTGAAGAGTCAAGCAGATCAGTTTGGCCAAAAGACTTGGGAAA	1213
QY	1201	tatgcgatgccttttgccttggactgcaatcccatltaatgaaatatgttagcagtgctbgy	1260
Db	1214	TATCGCATGCCCTTTGTGCTTGACTGCAATCCATTTAATGAATATTGTTAGCAGTGTGGG	1273
QY	1261	agtttgaaaagagatctcacaagaagtagaaatcagtaactgagaaagaaaggtcttgg	1320
Db	1274	AGTTTGAAAAGAGATTTCTACAGAAGTAGAAATCAGTACTGGAGAAACGAAAAAGGCTTGG	1333
QY	1321	tcagagagaggaattcctagtatgttgcgcagacgatacacttgaagacacaagaattgga	1380
Db	1334	TCAGAGAGAGGAATTTCTAGTATTTGTTGGCAGACAGATCACTTGAAGAGACACAGATGGA	1393
QY	1381	gatgatgcttgttaacttgcgcgagcctttcgaccagctactctcacagtgacaatttttt	1440
Db	1394	GATGATGCTTGTAACTTGACGAGCCTTTCGACCAGCTACTCTCACAGTGACAAATTTTTTT	1453

[illegible]

YDLLSLMDRGFEVNLIRHYCSQLSAKLSNLPFLISMRLERILCSHEHYLNLFEM
NADTAPTSPPSSISSONSSSSSFQDOKIASMFDLTSEYRQOHELTGLFTLAAALD
AEGEISKYORKAVSAIHSLSLSDHPVCVKPEVKVIAIYLPVGLITDALPOLC
DFTVADTRRYRTSGSDSEQEGAGAIQNONVALAIAGNNENLKTSGIVLSLPIKQYNL
NADTTRNLMLICFLWIMKNADQLIRKWIADLPSTQLNRLIDLEFCVLCEYKGOSS
DKVSTQVLOKSRDVKARLEALREBGARGEMRRRAPGNRFPGLNENLRWKKEQTH
WRQANEXIDKTKAEILDQELISGNLATEAHLIILDMQENITQASSALDCKDSLGLV
RLVNSLNCDSOTYYLTGCFATLRALIAKFGDLLFEEVEQCFDLCHQVLHCCSSMD
VTRSOACATLYILMRFSEGATSNPRVKMOTMSLASIVGARDNEEHLRSLRTIL
AYSEEDTAMOMTPPTQVEELLCLNLSILYTVKMRFEQDEMDLMYRIAKSYQA
SPDLRTWLOMAEKHTKKCYTEAMCIVHAALVAEYLSMEDHSYLPVGSVFQON
ISSNLVEESVSEDTLSPDEDEVCAGQYFTESGLVGLLEQAAELFSTGLEYETVNEY
KLIVIPILFAHREERKLTLTSHSKLQRAFDSIVNKDHRMFGEYRVEFGSKFGLDEQ
EFVYKEPATLTKLPEISHRLFAFYGCFAEFVEVIKIDSTPYDKTKLDPNKAYIQITEV
EPYFDEYEMKDRVTYPEKFNENLRREMYTPTFLEGRPRGELHQYRNTVLTMHAF
YIKTRISVIOKEEFVLTPIEVAIEEMKKTLQLAVAINOEPDAKMLQMLQGSVGAT
VNOGLEVAQVFLAEIRPADPKLYRHNKLRLCFKEFIMRCGEAVEKNKRLITADQREY
QOELKKNYNKLKENLRPMIERKIPLEYKPIFRVESQKRDSFHRSSFRKCEQTQLSQGS"

BASE COUNT 2005 a 1734 c 1721 g 1755 t
ORIGIN

Query Match 25.6%; Score 1749; DB 6; Length 7215;
Best Local Similarity 58.2%; Pred. No. 0;
Matches 3613; Conservative 0; Mismatches 2360; Indels 238; Gaps 21;

QY 173 ttaccgaagcagtagatccagtgatttgygaagattacatctactcatcttgctg 232
Db 71 TTTATGACCCCTGTGAGCCAGTGAGACTTTGAAGACTTGTATGACACACACCTGAACAGCC 130
QY 233 tggattctgagccttaccggaatttgatgaatttcctccagatgatattgaagtgtt 292
Db 131 TGGATGTGACAGCTTGCCACGAGCTCGGGACTTCACGTGATGACGACTTGGACGTGTGT 190
QY 293 atagtcctcgggactgcagaactctgttccagctgtacacctgaagaag--tgaatatg 349
Db 191 TCACGCCAAAGGAATGTAGGACTTTGCAGCCCTCTTTGCCGGAGGAAGGGGTGAACCTGG 250
QY 350 atccacatgttagagactgtataagaagltatacagaagactgygcaattgttcacagaa 409
Db 251 ACCCTCATGTGAGGACTGTGTTCAGACCTACATCCGTGAGTGGCTAATCGTGAACCGGA 310
QY 410 aatatcataaattggyaacagatttaaccatacatattagataaacagaagaagaagc 469
Db 311 AAAACCAAGGAAGTCCAGAAATCTGTGGCTTTAAAAAGACTGGATCTCGAAAGATTTTTC 370
QY 470 aaaaagtttgccaaaacaagtlttgaatcgtatgaagctccagatgyccaacagctacc 529
Db 371 ACAAGACGCTTCCGAACACAGCTTTGAGTCGGA-----AACCTTGGAGTGCAGTGAAC 424
QY 530 aggatgatcaagatgaccttaaaaagcgttcaatgtcaatagatgataccccaaggggta 589
Db 425 CCGCTGCTCAGGCAAGGCCCCGCCACTTAAACGTGCTGTGCGACGTGTCTGGGAAAGGCC 484
QY 590 gctgagcctgtagatatttgaacttgaaaaattcacttccctgatgtcttgcctccaatt 649
Db 485 CCGTCACTGCTGTGACTTTGACTTCGCGACGCTGACGCTGACCAAGCGGCTAGAAAAACC 544
QY 650 tacttgatcgaaactccaatgaagaataagacctcagaatgatgaccaagaagaatcaa 709
Db 545 TCCTGCGACCAAGTGAGTGCAGGACTTTGAGAAAGACGAGAGAGGCCCGGAGGACCA 604
QY 710 accgtcaaaaagaacttttgccttgcattcatcatcaccagatgaggaagaaccaatagaac 769
Db 605 ACAGGCAAGCCGAGCTCTTTGCCCTTTAACCATCAGTGGACGAGAGAGATGCTGTGAAAA 664
QY 770 ggcttagtgttccctgatataaccaaaagaacatttgttcaaaagacttcttgaataatgct 829
Db 665 TACGTCCAGTACCAGAATGTCCCAAGGAACACCTGGGCAACAGAAATATTGGTCAAGTTGC 724
QY 830 tatcaactcaagtttgaattgaaccatttltgcaagtttggcttlatatgatg 889
Db 725 TGACCTTGAAGTTGAGATTGAAATTGAGCCCCCTGTGTGCGACGATTTGCCCTCTACGATG 784

QY 890 tcaaggaanaaagaanaagatttcagaanaactttatttgcacttaattctgagcagatga 949
Db 785 TTAAGAAAGGAAGAAAAGATCTCAGAAAATTTTCACCTGTGACCTGGAACCTGACCCAGTTCA 844
QY 950 aaggtgtttagcgtccacatgtaccacactgtctgccattactacccttgcagaagatcagcaa 1009
Db 845 AAGGATTTCTGCGAGCTCACACCGCCTTTCAGTGGCCGCAATCAAGTCAGGCGAGATCTGCAG 904
QY 1010 ttttctcatcactatccctcccaagaatgttcttctgttaataaagctagaanaagctcc 1069
Db 905 TCTTCTCAGTCACTACCTACCCGCTCTCAGACATCTACCTGGTAGTCAAGATTGAAAAAGTCC 964
QY 1070 tacagcaaggaagacattgagagatgtgcagaaccatatagtatltcaagaagcagatg 1129
Db 965 TGCAGCAGGAGATATTGGAGACTGTGCAGAGCCCTACACGGTTATCAAGAAAGTGTATG 1024
QY 1130 ccaccaagaataaagaanaaactggaagaactgaaagatcaagcagatcagtttgcacaa 1189
Db 1025 GTGGAAGAGTAAGAAAGATTTGAATAAACTAAACCTCAAGCTGAATCCTTCTGCCAGC 1084
QY 1190 gacttgggaataatcgcatactgtccttctgttgcactgtcgaatccatttaatatgttta 1249
Db 1085 GTTGGGGAATAACCGGATGCCCCCTTGCTGGGCAACCCATTAAGCTTACTTCAAGCTTCT-- 1142
QY 1250 gcaatgcttgggaagtttggaaagagatltcacagaagttagaatacagtlactgagagaagaa 1309
Db 1143 -CAATGTCCTCACCCCTTGAGAGAGGAGTAACCTGATGTGGA-----CTCTGTGCTTGGGA 1195
QY 1310 aaggtcttgtcagagagagaagaattctaglatgttgcagacagatcacttgaagaaga 1369
Db 1196 GAAGCCAGTGGGTGAACGAGGACATTTGCCCAATCTAGAAGGCTTTTC--TGAAGAG 1252
QY 1370 caacaagtggagatgatgtgttaacttgaagcgttctgcagcagctactctcacagtga 1429
Db 1253 CCCTCTCCTTGAAGAAATAGGGTTGGATCCAACTTCAAAACCTCCACTCTGAGCGTTA 1312
QY 1430 caaattttttaaagcagaagagagaccgcttaagtgtatgaagatctctacaattcccttg 1489
Db 1313 GCAGCTTTTTCAGACGAGAGAGAGATGCGCTTAGCGATGAAGACTTATTCAGATTTTTAA 1372
QY 1490 ctgatatgagaagggcactctctgtctttaaagcgactaagacattatcacagctcagctca 1549
Db 1373 CTGACTTACAAAGATCATCATCTCTTACAGAGACGAGTCAAGTCAATTCCAGGCTTGCTTAA 1432
QY 1550 agatagacattctcccgacactgaanaatcccatatttgcctaactccggagctgtcc 1609
Db 1433 GACTGGAGATTCTACAGCTCCAGAGATCATCAATTGCTGTGCACTCCTGAATGCTGC 1492
QY 1610 aagtgaagctttaccctgacagtagagttagacctaccagagaanaatcttagagttcccg 1669
Db 1493 CCGTGAACCCCTTCTCTGAAAAACCGGACACGCCCGCACAAAGAGATTTTGAATTTCCAA 1552
QY 1670 caagygatgttatgttccaacacttacttacagaanaatcttctctacatataccctcaga 1729
Db 1553 CACGAGAAGTATATGTCCCTCACACTGTGTACAGAAACCTTCTCTATGTCTACCCACAGAGA 1612
QY 1730 gtcttaatttgcacaatcgtcaaggttctgtctagagaataataacagtgaaagtccaagtta 1789
Db 1613 GGCTGAACCTTTGTAAACAAACTAGCATGAGCCCGGAAACATTACAAATAAAGATCCAGTTTA 1672
QY 1790 tgtatgagagatcccaagcaatggcagtcgggtaatcttgtgtaaatctagctgttcag 1849
Db 1673 TGTGTGAGAGAGATGCTAGCAATGCGATGCCGCTCATCTTTGGAAAAATCCAGCGGCGCTG 1732
QY 1850 aatttcaagaagcctatacagccgtagtatatacataacaggtctcctgatlttcag 1909
Db 1733. AATTCTGACGAGAGTGTACACAGCTGTATTACATACCATTAATTAAGTCTCTGACTTTTATG 1792
QY 1910 aagaatcaaggttaagcttccctgtacttlaactgaccatcatcactgtcttcttact 1969
Db 1793 AAGAAGTGAATAATTAAAGCTCCCGCTAAGCTCACAGTAATAATCACACCTCCTGTCACT 1852

QY	1970	ttatcatgttagttgttcaacaaacaaataactcctcttgaacacacagttgatat	2029
Db	1853	TCTACCATATCAGCTGTCAACGAGAAGAGAGCCTCCGTGAAACTCTCCTGGATATT	1912
QY	2030	catgataccaatgcttcagaatgagcgtttgaagactggccagtttgtctgcagttc	2089
Db	1913	CATGGCTGCCAATTCTTTAAATGAACGCTTCAAACTGGATCCTACTGTCTCCAGTTG	1972
QY	2090	cattgaaaaaacccacacagccttatctgtactgtctcctgag-----gttcctc	2140
Db	1973	CCTTGAAAAATTGGCCACCCTCACTACTCCATGCATTGTGCTGAGAAAGTCCCATTA	2032
QY	2141	tacctgcatgaatggtgtagataatacacaagtgtrtttaatgttgaagttgtgtg	2200
Db	2033	ATCCTCCCATTAAGTGGGCTGAAGACATTAAGGAGATTTAATATTGAAGTGCACCTG	2092
QY	2201	ttcgtctatccatacacaagatccctatccttgacaatttttgcctcgttcaatgctc	2260
Db	2093	TTTCTTCTGTACACACCACGACCAACACCTGGAGAGTTCTTACCCCTCTGCCACTCC	2152
QY	2261	tgatatgaacacgt--ttcccagtcogaattgggacatgcyaatcatgtgaaataact	2317
Db	2153	TGGAGAGCCAGGTGACCTTCCCATCCGCTGCTGGATCAGAAAAATCAGCGAGATGGCGC	2212
QY	2318	tagaaaatgaattgaagagcaglattcagcacgtgaatttccacagctggaaccagtg	2377
Db	2213	TGGAGCATGAGCTGAAGCTCAGCATCATCTGCTGTAATCCTCCCGCTGGAGCGGCTCG	2272
QY	2378	tccgatttcttcatctctcgtagataaactgtaactttagttattagacctcctgtca	2437
Db	2273	TGCTCTTCTTGACACCTGGTGTGTTGACAAAGCTCTTCCAGCTGTCCGTGCAGCCCAT	2332
QY	2438	ttgtcggccaatagttaaacctaggtcagaacatcttttgaagccatgycatcaattata	2497
Db	2333	TGCTGGCCAGACAGCCCACTTCTCCAGTTTGCCCTTGAGTCCGTGGTGGCCATCGCCA	2392
QY	2498	atcgacttcacaaaaacttggaaagaaatcatgaccagcatgycagaacacagcctctg	2557
Db	2393	ACAGTCTGCACACACAGCAAGAACCTGAGCAAGGACCAAGCATGGAGGAACCTGCTGG	2452
QY	2558	calcatatatcatltaatgttttcgcctaccacaacttacccta--tcatcatcac	2614
Db	2453	CTTCTACGTGACACTACGCTCTTCGCGCTGCCAGAGGTGCAAAAGGATGTGCCAAGTCA	2512
QY	2615	caggtcctggggttttgggagagatcagtgcatatgccaacatgycatagatctgcgtga	2674
Db	2513	GCGCTCCCACTGCCCTCTTAGACCCCTCGAGCTACCAACAGTATGGCCGACATCAGCTG	2572
QY	2675	gacctgcaagccttaatttaaatcgltctcgaagccttagtaataagcaatccagatat	2734
Db	2573	CTGCTGTGAGTTCAAAAGCTGTGTCAGGCGCGGTTGATGAGCAGCAGTAACCCAGACCTCG	2632
QY	2735	ctgagactcccaagtcacaccagatgataagttcgaatcaatcatcgcggagtaagggttag	2794
Db	2633	CGGGGACACACTCCGAGCAGCAGAGGAGTAGAAGACATCATGTCTTCAAGATCGCCG	2692
QY	2795	atcgctccaattccttgggttaacactggtgttccaataaagctgccccatggygatcca	2854
Db	2693	ATCGCAACTGCAGC-----	2706
QY	2855	ccagtccaagtcagaatacaacacagctatgatacgaagttgtaatcgatgtcctgc	2914
Db	2707	-----CGAATGTCTTACTATTGCTCTGGCA	2731
QY	2915	acacagagacgtcaagttcttacaacatbaacggyagccttaccaactaaagaactt	2974
Db	2732	GTAGTGATGCTCCAAGTT-----CACCTGCAGCCCCAAGGCCAGCCAGCAAAAAGCATT	2785
QY	2975	ttcacgagagcgttggtcttgcagttggttgtttgcagtgycagcggttcggyaatcagct	3034
Db	2786	TCCATGAGGAGCTTGCCCTTCAGATGTTGTTCAAGCACCG-----	2825
QY	3035	tgcaacaagcctgttcttttttgaatlaatgyltaagagcatggtgcaccattatact	3094

D	b	2826	-----AATGGTGA	AAGCATGCGCCACGACCGTACATA	2857
Q	y	3095	ttaatgataaacttgaaggtcccaaggaaaagtcgtttccagaacglttcatgatgaca	3154	
D	b	2858	ACATGGACAACGGGACAGTTTCGGAGAGCTCGTTTTTGACCGGTTCATGATGATGACA	2917	
Q	y	3155	ttgcagctctgtcagcaacgattgtctaigtalatagtltcacgatt-----tcaga	3205	
D	b	2918	TAACTACTATTGTTAATGTGTCACCCTCGAAAATTGCAGGCCCTTTTAGTAAMAACACAGA	2977	
Q	y	3206	aggacacaganaatggtltgagacctcaatacaaacgcttgcatcttcctcaatgatctgt	3265	
D	b	2978	AGGAAATGAACAGCGGAAAAAGATGACATCAGCCTGGCTTCTTCTGTATGACCTTC	3037	
Q	y	3266	tgtctgttatgtagcagaggatttgtttttaagcctlataaagtcctgtctataacaggtgt	3325	
D	b	3038	TCTCCCTCATGGAATCGGGGCTTGTGTATTAACTCATCAGACATTATTGCAGCCAGCTGT	3097	
Q	y	3326	cttcaaagcttlactcatlaccgnaatcccagtgltctgtgtctccttgagcttgatlctc	3385	
D	b	3098	CAGCCAAGCTCAGTAACCTTCCAACGCTCA-----TTTCCATGAGGCTAGAGTTCC	3148	
Q	y	3386	tacgaatcatctgcagtcagtcatgagcaactatgtlacaattaacctaaccctgcagctta	3445	
D	b	3149	TGAGAACTCTGTGTAGCCATGAGCAATTACTCAATCTGAACCT----TTTATTATGA	3202	
Q	y	3446	ctccacctgcacatccaatcacctctcgtttcctctgcacaacatccagagttctgatttc	3505	
D	b	3203	ATGCTGATACCTGCTCCAACATCTCCTGTCTCTTCCATATCTTCCCAGAACTCAAGCTCT	3262	
Q	y	3506	ctacgaatgtlaaaccaaaagattgcaaatatgtttgaattatccgtgccttcgcc	3565	
D	b	3263	GCTCCAGCTTCCAGAGACCGAAGATGCCACGATGTTGCATCTGACTCCGAGTACC GCC	3322	
Q	y	3566	aacagcatatttggcaggaacttgtttaaagagctgtgcgtgtcatltagaccctgatg	3625	
D	b	3323	AGCAGCACCTCCCTCACCGGGCTCCCTCTTCACAGAACTGGCTGCTGCCCTGATGCCGAAG	3382	
Q	y	3626	ctgaaggaactgttttgatttgcatataaagaagtcataatatgytlacacaaatlactlcca	3685	
D	b	3383	GGGAAGGAATCAGCAAAGTACAAAGAAAGCTGTCAGTGCAAATTCCACAGCCCTGCTAAGTT	3442	
Q	y	3686	gtcacgactcagaccgcggtactctgacccctcagataaagctcgagtgccatgtgt	3745	
D	b	3443	CTCAGACCTGGACCCACGCTGTGTCAAAACGAGAGGTGAAGGTCAAAATCGCCGCTTT	3502	
Q	y	3746	atctaccttgattggtattatcatgysaaactgtacctcagctgtlatgatltaacagaaa	3805	
D	b	3503	ACCTACCTTTAGTTGGCATCATTTTGATGCTTTGGCACAGCTCTGTGACTTTACAGTTWG	3562	
Q	y	3806	ctcacaaatcaacgaggaagaccaattgtatagccactgatagtatgaagytagagcg	3865	
D	b	3563	CAGATACTCGCAGATACCGCACAGTGGCTCG-----GATGAAGAACAAAGAGAG	3613	
Q	y	3866	gaagtatgatalaagccagaccgtltgcoatgycaatgcgagggacatcggtccctcaactaa	3925	
D	b	3614	CCGGTGCCATTAAACAGAAATGTGGCTCTGGCCATTAGCAGGGAATAATTTCALTTTGAAAA	3673	
Q	y	3926	caaggcctgycagtlctcctcctcaagtcaacgagtgycagggcaacacactactclttcag	3985	
D	b	3674	CAAGTG-----GAATAGTGTGTCTTCTTCCCTTGCCCTATAGCAGTACAAACATGCTGAACG	3727	
Q	y	3986	cagaatcaagtcgaagccctltgactgtctactlttgygttccaanaatgcagatgaaa	4045	
D	b	3728	CGGACACTACTCGCAACCTCATGATCTGCTTCTCTGTGATCATGAAAAATGCTGATCAGA	3787	
Q	y	4046	cagttctacagaagtggtttacagatctctcagttctlgcagctaaaccgycattatgac	4105	
D	b	3788	GCCTCATTTAGGAAGTGGATGTGTGACCTGCCATCAACCGACGCTCAACAGGATTTTAGATC	3847	
Q	y	4106	tgctttatcctctgtgtctctgtcttgagtalataaagggaaaaaagtglttgaaagaaatga	4165	

Db 3848 TACTTTTCATCTGTGTATGTTTGAGTATTAAGGAAACAGAGTTC TGACAAAGTCA 3907

QY 4166 atagcttgacctttaagaaatcaaaaagacatgagagcaaaagcttgaagaagtattcttg 4225

Db 3908 GTACCCCAAGTCCCTGCAGAGATCAAGGGATGTCAAGGCCCGGCTGGAAGAGGCTTTGCTGC 3967

QY 4226 ggaagcatagtgccagcaagaaatgylacgycgaagccgagjacagctcgaagaaagcc 4285

Db 3968 GTGGGGAAGGGGCCAGAGGGAGATGATGCGCGCGGCTCCAGGGAACGACCGATTTC 4027

QY 4286 catctggaagtgccttgggaagtcaagaaattgagtggaagaaagatatgactcact 4345

Db 4028 CA-----GGCCTAAATGAAAAATTGGAGATGGAAGAAAGAGCAGACACATT 4072

QY 4346 ggcgtcaaaacacacagagaagcttgacaatatcaagagcagagattgaaacagaaagcactga 4405

Db 4073 GGGGGCAAGCTAATGAGAGAGCTAGATAAAACAAAGCGCAGTTAGATCAAGAACCTTGA 4132

QY 4406 ttgatggaaacctggtctacagaagcaaacctaattcatlttagatatacattagagattgtg 4465

Db 4133 TCAGTGGCAATCTGGCTACAGAAACACATTTAATCATCTGGATATGCAAGGAAACATTTA 4192

QY 4466 ttcagaaccgttctctgtaacggaatccaaagagacatctctgtgtgagtgctlaaagtgc 4525

Db 4193 TCCAGGCCGAGCTCGGCTCTGGAAGTAAAGACAGCCTGCTGGAGGTTGTTCTGAGGGTGC 4252

QY 4526 tactacaagcaatgycgtgtlaaccaaagtgcaatttatactacaacactgtttgtctaac 4585

Db 4253 TGGTGAATTCCTGAACTGTGATCAGAGTACCACCTACTGACTGCTTTCGAACAC 4312

QY 4586 agaagacctgtgttcaaaagtttcctgaactcttatitgaaagaagagacagagcagtgtg 4645

Db 4313 TCCGTGCTCTCATCGCCAAAGTTGGAGACTTACTCTTCGAAGAGAGGTGGAACAGTGT 4372

QY 4646 ctgatttatgcctcagagctctccgacactgtagcagtagcatcgylataacaggtcac 4705

Db 4373 TCGACCTATGTCACCAAGTCCGTGCACCACTGCAGCAGCAGCATGGATGTCAACCCGAGCC 4432

QY 4706 accccaagtccctccttacctactaatgagggcaaaactttgagattgggaataactttg 4765

Db 4433 AAGCCTGTGCCACCCCTTACCCTCCTCATGAGGTTCAAGTTTGGAGCCACCAAGTAAATTTTG 4492

QY 4766 ccaaggttaaaatgcaagtlaccaatgltcaatcctcctcgtgtgycacatcagaatt 4825

Db 4493 CAAGAGTAAAGATGCAAGTAAACCATGTCCCTGGCATCTTGGTGGAAGAGCACAGACT 4552

QY 4826 ttaatgaagaattcctaagacgttctctaagaactataatgacatalgtcgaagaagatc 4885

Db 4553 TTAATGAAAGACACCTGAGAAAGATCCTTGAGGACAATTGGCCTATTTCAGAAAGAGACA 4612

QY 4886 tgyaatgaggggaaacaacatttcctgtatcaggtccagatctgthttcaatctccata 4945

Db 4613 CAGCCATGCAGATGACTCCTTTTCCACCCAGGTGAGGAACTTCTCTGTAATCTGATA 4672

QY 4946 tgaattcttctgtactgtgaaaaatgaaagaaacaccagagagatcctgaaatgttgattg 5005

Db 4673 GCATCTTAATGACACAGTGAATAAGGGAAATTTCAGGAAGATCCTGAGATGCTTATGG 4732

QY 5006 atctaagtacagaattgccaaggtttaccagaacctccaga--gcgattgacctggt 5062

Db 4733 ATCTCATGTACAGAAATTGCCAAGGTTACCAGGCATCTCTGATCTCGGCTGACCTGGC 4792

QY 5063 tgcagaacatgycagcaagcactcaagaacgaagcaatcatgctgnaagctgcacaggtgc 5122

Db 4793 TCCAGAACTGGCAGAGAAACACCAAGAAGAGTGTACACGGAGGCTGCCATGTGCC 4852

QY 5123 tagtccactcagcagcactgtgtgtgaatatlttgacatgctggagaccggaataatc 5182

Db 4853 TGGTGCACGCCGCTGGCTTAGTGCTGAGTATCTGAGCATGCTGAGAGCACCACAGCTAAC 4912

QY 5183 ttccgtgagatgtglaacattccagaataatltcatctaatgtlttagaagaatctgcgg 5242

Db 4913 TGCCCGTGGCAGTGTACAGCTTCAGATAATTTCTTCCAAATGTGCTGGAGGAGTCTGTGG 4972

QY 5243 tctcagatgatgtgtatctccagatgaagaaggtatctgctctlgnaaatacttactg 5302

Db 4973 TCTCTGAGGACACCCCTGTCAACCTGACGAGATGGGTGTGCGCAGGCCAGTACTTCACCG 5032

QY 5303 agtcaggaactgttggattactgtgaacaagcagctgtcttcctctctatgtgctgcagt 5362

Db 5033 AGAGTGGCCTGGTAGGCCCTCCTGGACGAGCGCGGAGGCTCTTCAGCACGGGAGGCTTAT 5092

QY 5363 atgaagcagttaatgaagtttacaagtlacttattcctatcatgaagctaatcggagtg 5422

Db 5093 ATGAGACAGTTAATGAGGTCTACAGCTGGTCATCCCATCTTAGAAGCGCATCGAGAA 5152

QY 5423 caaagaactataccaatltcatgttaaacttcaagaagcattcagcaaaatgttcatc 5482

Db 5153 TCCGGAAGCTGACACTCACTCACAGCAAGCTGCAGAGAGGCTTCGACAGCATCGTTACA 5212

QY 5483 agagtactgctgygagcagatgttgycaactalttgcgtgtgtgtlttatgaaacca 5542

Db 5213 AGGATCAT-----AAGAGAAATGTTGGAACCTACTTCCGAGTGGTTCTTTGGATCCA 5266

QY 5543 agttcgggatttggatgaaacaagaattgtlttaacaagagcctgcaataaccaacttg 5602

Db 5267 AATTGGGGATTGGATGAACAGAGTTTGTCTACAAAGAGCCTGCATATTACCAAGCTTC 5326

QY 5603 cagagatatctcacagatttggaggttltacygagaagaatttggagagatgtgtgtg 5662

Db 5327 CTGAGATCTCACATAGACTAGAGGCATTTTATGTCATGTTTTGGTGACAGAAATTTGTGG 5386

QY 5663 aagtaatcaaaagactctaactcctgttagacaaagltglaaattagatcctaacaagcadata 5722

Db 5387 AAGTGAATTAAGACTCCACTCCTGTGACAAAACCAAGTTGGATTCCTAACAAAGGCTTACA 5446

QY 5723 ttcagatlacctatgtlygagccatacttgaacacatalgagatgaaagacagaatacact 5782

Db 5447 TACAGATCACTTTTGTGGAGCCCTACTTTGATGAGTATGAGATGAAGAAGACAGGGTCAAT 5506

QY 5783 atttcgacaaaattacaatcttgcgtacattcaatgtlactgtlacaacccttactttagtg 5842

Db 5507 ACTTTGAGAAAGATTCAACCTCCGGAGGTTCATGTACACACCACCCGTTCAACCCGTGAGG 5566

QY 5843 gccgtgcccatggyggaactltcatgnaacaatlcaaaagaagaccaatctgactagctc 5902

Db 5567 GGGGCGCTCGGGGAGAGCTGCATGAGCAGTACAGAAGAAACACAGTCCGTGACCACTATGC 5626

QY 5903 atgccttccattataattaacaaggytcaatgtlcaactcataaagaagaagatcatcttaa 5962

Db 5627 ACGCCTTCCCTTACATCAAGACCAAGATCAGCGTTCATCCAGAAGSAGGAGTTGTTTTGA 5686

QY 5963 caaccaatlgaagtgctalttgaagacatgcagaaaaagacacaggaagltgycattgcca 6022

Db 5687 CACCGATTGAAGTGGCCATTGMAAGACATGAAGAAAGAACCCCTGCAGTTAGCAGTTGCCA 5746

QY 6023 cacatcagatcccgcagaccccaaatgcttcagatggtlactccagygatcgttagca 6082

Db 5747 TTAACCAAGAGCCGCTGATGCAAAAGATGCTTCAGATGGTGTGCAAGGCTCTGTGGAG 5806

QY 6083 ccacagtgaaatcaggggoccttggaaagtgcocaggtlttctgtctgaatatcattagt 6142

Db 5807 CTACTGTAATTCAGGGAGCACACTGGAAGTACCCCAAGTGTTTTGGCTGAATTCCTGCTG 5866

QY 6143 acccaagctcttcagacatacataataaactgcactctgctttaaagaatttactaaaa 6202

Db 5867 ATCCAAAACCTCTATTCGACATCACACACAGTTGAGGTTATGTCTTTAAGGAATTTCATCA 5926

QY 6203 ggtgtgaagatgccttaagaaaaaataaagccttaattggyccggttcaaaagagatc 6262

Db 5927 GATGTGTGAAGCTGTAGAGAAAAACAACGCTCTCATCAGCGGACAGACCAGAGGAATATC 5986

QY 6263 aaaggaattgggga--aactatctgcctttaaagagccctacagccctagatc--ac 6318

Db 5987 AGCAGGAACCTCAAAAAGAACTATTAACAAGCTAAAAAGAGAACCTCAGGCCAATGATCGAGC 6046

Db	905	TCGACCTGAAC	TCGGACTCCAT	TGAAGGGGGCTTCGGGCTCATGGACACCCCTGCCA	964
QY	986	ttactaccctg	caagatcag	caatttctctatcacttacccttcccaagatgttttc	1045
Db	965	TCTCCACCCT	TGGCCCGCT	TGCCATCTTCTGTGACCTACCCCTCACCCTGACATCTTCC	1024
QY	1046	ttgtataaagt	tagaaaaagtc	ctacagcaaggagacattgagagtggtgcagaacct	1105
Db	1025	TGGTCATCAAG	TGTGGAGAGGT	GCTTTCAGCAAGGGGACATCAGTGA	1084
QY	1106	atatgat	ttccaagaagcagatg	ccaccagaataaagaanaactgagaaaaactgaaga	1165
Db	1085	ACATG	TGTGAAGAAGT	GGACACAGCCCAAGAACAAAGAAGTAGAAGCTGCGCC	1144
QY	1166	gtcaagcagat	cagtttgc	caagactggygaatatcgcatgcttcttgactg	1225
Db	1145	TGGCGGCCG	AGCAGTTCGAC	ACCGCCTGGGGCCTACCGCATGCCCTTGCGCTGGACGG	1204
QY	1226	caatccatt	taatgaatat	tgtagcagtgctgygagtttgyaaagagatctacagaag	1285
Db	1205	CCGTGCAC	TGGCCCAACAT	CGTGAGCAGCCGTGGGCACTGGACCGGACTCTGACTCGG	1264
QY	1286	tagaatacag	tactgagaa	cgyaaagtgcttgcagagaggaattctaglatg	1345
Db	1265	AG-----	GGCGAGCGCGCG	CAGCCTGGACAGACCGCGG-----	1298
QY	1346	ttgycagacg	atcactgaa	agacacaagtgagatgatgtgttaacttgcagagct	1405
Db	1299	-CGTCGGGG	CGCCAGACCGGG	CGAGTAGTGGGGACGACGCTGCAGCTTCTTGCT	1357
QY	1406	ttcgaccag	tactctc	acagtgacaatttttaagcaggaagagacgcgttaagtg	1465
Db	1358	TCCGTCCAG	CCAGCGTAACT	GTCAACAACCTCTTAAGCAGAGGCGTGAGCGACTCAGTG	1417
QY	1466	atgaagat	ctctaca	aattcccttgctgatatgagaagccatcttctgtcttaagcgac	1525
Db	1418	ACGAGGAC	CTCTTCAAGT	TCTTGCGTGACATGAGGCGCCGTCGTCCTGCTGCGCGAC	1477
QY	1526	taagacct	tattacag	ctcaagctcaagatattctccgcacctgaaaaatccccatt	1585
Db	1478	TACGTCTGT	GACTGCCAC	AGCTCAAGATTTCTCCGGCTCCTGAANAATCCCCACT	1537
QY	1586	attgccta	actccgcg	agctgtctcaagcttaaccctgcagtagagttagaccta	1645
Db	1538	TCTGCC	TCTCCCCGTG	AGCTGCTCATATCAAGCCCTAACCCGGACCCCAAGGGCGGCCCA	1597
QY	1646	ccagaga	aatccttag	agttcccgcaagygatgttatgttccaacactacttacagaa	1705
Db	1598	CCAAGGAG	ATTCTGAG	ATTCCCCCGCGGAAGTCTATGCCCCCATACCAAGCTTACAGGA	1657
QY	1706	atcttct	catata	taccctcagagtccttaatttggccaatcgtaaggtctgctagaa	1765
Db	1658	ACCTGCTG	TACGTATACC	CGCACAGCCTCAACTTCAGCGACCCGCAAGGGCTCCGTGCGCA	1717
QY	1766	ata	taacag	tgtaaaagtcagtttatgtlatgagagagatccaagcaatgccaatgcggtaa	1825
Db	1718	ACCTTGC	TGTGCGAG	TGCAATGACAGGCGGAGGACCCCAAGGCTCTGCGGTCA	1777
QY	1826	tcctt	gtaaat	ctagctgttccagaatttccaagaagcctatacagccgtagtatac	1885
Db	1778	TCTTTGG	CAAGTCCAG	CTGCAGTGAATTTACCCGCGAGGCGCTTACACCCGGTGTACCC	1837
QY	1886	ata	caagctc	ctcctgatatttcaatgaagaatcaaggttaagcttccctgctactttaa	1945
Db	1838	ATAACAAG	TCCCCCGAG	TCTTACGAGAGTTCAAGCTGCATCTTCCAGCCTGCGTGACAG	1897
QY	1946	acc	atcatc	acttgcttttacttttatcatggttagtgccaacaaaacaaatac	2005
Db	1898	AGAACC	ATCACCTG	CTGTACCTTCTACCATGTCAAGCTGCACGCCCGCGGCGCACTG	1957
QY	2006	ctcttg	aaacaccag	gttgatatatacatgataccaatgcttcagaatgacaggttgaga	2065

Db	1958	CCCTGGAGACACCCGTTGGCTTTACTTTGGATCCCACTGCTGACAGCACAGGCGCCTGAGGA	2017
QY	2066	ctggccaglttctgtctgcccagttcatlttgaaaaaacccacaggttattctgtactgt	2125
Db	2018	CGGCGCCCTTCTGTCTCCAGTGTCTGTGGACCAGCGCGCGCCAGCACTATTCGGTGCTCA	2077
QY	2126	ctccctgaggttcctctacacctgycagtgaaatggytagataatacacaaagylgtltaatg	2185
Db	2078	CACCCGATGTGGCGCTTCCGGGCAATGCGCTGGGTGGACGGTCAACAGGGCGTGTTCAGTG	2137
QY	2186	ttgaagttgtgtctgttctgtctatcatcacacaagaatcctatcttgacaaattlttg	2245
Db	2138	TGGAGCTCACGGCCGTGTCTCTGTGCACCCCCAGAGACCCCTACCTGGACAATTTCTCA	2197
QY	2246	ctctgtgtaaatgtctctgcatgaaacacctgttcccagtcggaattggygacatgcyaatca	2305
Db	2198	CCCTGGTGCACGTCCTTGAGGAGGAGGCCCTTCCCATTTCCGGCTCAAGGACACTGTGTGA	2257
QY	2306	tggaaaaataacttagaaaaatgaaatgaaagacaglatttcagcaactgaatttcatccagc	2365
Db	2258	GCGAGGGCAACGTGGAGCAGGAGCTGCGGGCCAGTCTTGACGACTGCGCCTGGCCAGCC	2317
QY	2366	tggaaaccagtggtccgatttctcatcttctgtagataaactgatactttagtlatla	2425
Db	2318	CCGAACCCCTTGTGGCCTTCTCCACACGCTGCTGGACAAAGCTCGTGGCTGTGTCATCA	2377
QY	2426	gacctcctgtcattgtctggyccaaatagttaaacctaggtcaagcatctttgaaagccatgy	2485
Db	2378	GGCCCCGATCATCAGTGGCCAGATGTGTGAACCTGGGCGGTGAGCCCTTGAAGCAATGG	2437
QY	2486	catcaattataaatcgacttcacaaaaacttgyaaagaaatcatgaccagcatgycagaa	2545
Db	2438	CCCATGTAGTCAGCCTTGTTCACCCGGAGCCTGGAGGCCAACCAAGATGCCCGCGGTCACT	2497
QY	2546	acagcctcttgcatacatataatcatltagtlttccgcctaccaaatacttaacctaat	2605
Db	2498	GCCCAAGCTGGCTGCTTACGTCCACTACGCTTTCGCCCTTCCTGGCACTGAAGCCCA---	2554
QY	2606	catcatcaccagglcctggyggttgygagatcagltgcataatgcccacaatggtctagat	2665
Db	2555	-----GCCCTCCCGGATGGGGCCCTCCAGTAGACAGTGCAGGCTGCCACACTGGCCCGTG	2608
QY	2666	ctgcytgagacctgcaagccttaattaaatcgttctcgaagcctagtaataagcaatc	2725
Db	2609	GCTCTGTCGCCCCCGCAAGCCTTACCTGGCGGCTTCCAAAGACATCAGCAGCAACAAC	2668
QY	2726	cagatatatctgygactcccaagylcacccagatgtaagttcgatcaatcatcogggagta	2785
Db	2669	CTGACCTCGCCGTGGCCCTTGCTCTGTGTGATGACGAGGTTTCCCGCATCCTGGCCAG--	2726
QY	2786	agggttagatcgctccaattcctggttaacactggtgtccaaaagctgcccacatgyg	2845
Db	2727	-----	2726
QY	2846	gatccaacccagltccaagltgcagaatcaacacaggtctatygatcgaagltgtaatcgta	2905
Db	2727	-----	2726
QY	2906	tgtcttcgacacagagacgtcaaglttcttaacaacaltaacgagccttaaccaacta	2965
Db	2727	-----	2726
QY	2966	aaaagcttttcacgagagagctgagcttgcagfgyggtgtgtgcagtgycagcgttcggy	3025
Db	2727	-CAAGCTGCTTCACGAGAGAGCTGGCTCTGCAGTGGGTGTGCAGCAGCAGTGCCTGACGG	2785
QY	3026	aatcagcttgcacaacaagcctgttcttlttgaaatgaatggttaaagacatggtgcacc	3085
Db	2786	AGGCCATCTCCAGCACCGCCTGTTCTTCTTCACGCTCATATGGTGAAGATATGGCGCTGC	2845
QY	3086	attataactttaatgataaaacttgaggtcccaagaaaaagtcgttctccagaagctttca	3145
Db	2846	ACCTGCTGCTTGGCCAGCGACTAGACACACACCCGCAAGCTTGCGCTTCCCGGAGCGCTTCC	2905

QY	3146	tgagatgacatttgcagctctgttcagcagcagattgtctaagtataagtttccagattccaga	3205
Db	2906	TGGACGACATCACTGCGCTTGGTGGGCTCTGTGGGCGCTGGAGGTCAATCACCCGTGTCCACA	2965
QY	3206	aggacacagaagaatggtttgagagacctaatacagcccttgcattcttctcaatgatctgt	3265
Db	2966	AGGATGTGAGCTGGCGCGAGACCTCAACGCCAGCCGTGGCTTCTTCTCAGTGACCTTC	3025
QY	3266	tgctcgttatgagacagagattgtttttagccctataaagtcctcgtcctataacagtggt	3325
Db	3026	TGTCCCTGTGTGACCGGGCTTGTCTTCAGCCTGTGTCCGGGCCACTACAGCAGGTGG	3085
QY	3326	cttcaaaagcttactcatctaacccaatcccagtggtctgtgtctcctttagagctgatttc	3385
Db	3086	CCACGGCGCTCCAGTGTGTCCCGCTAATCCAGCAGCCCTGTGTGAACCTGCGCATGGATTCA	3145
QY	3386	tacgaatcatctgcagtcacatgagcactatgtttacatlaaacttaacctgcagcttaacta	3445
Db	3146	CCCCGATCCTGTGTGACGCCACGAGCAGCACTACGTGACCCCTCAACCTCCCTGTGCCCCCTGT	3205
QY	3446	ctccacctgcatactccatacccttctgttctctctgcaacatctcagaagttcctgattt	3505
Db	3206	CACCTCCAGCCTCGCCCTCCCTCTGTGTCTTCACCACTCCACAGACTCCACCTTCT	3265
QY	3506	ctacgaatgtacaagacccaagaatttgcaaatatgtttgaattatccgtgccttcgcgc	3565
Db	3266	CCAGCCAAAGCCCCGAGCCCCCAAGGTGACCAGCATGTTCGAACCTGAGTGACCATTCGGC	3325
QY	3566	aacagcattatttgcgcaggacttgttatacagagcctgcgtcattttagaccctgatg	3625
Db	3326	AGCAGCACTTCCTAGCTGGGCTCTCTGTGACGGAGCTGGCACTGGCCCTCGAACCTGAGG	3385
QY	3626	ctgaaggactgtttggaattgcataaagaagtcatacaatatgttacacaatttactctca	3685
Db	3386	CTGAAGGGGCAATTCCTGTGTGCACAAGAAGGCCATCAGTGTGTGCACAGCCTGTATGTG	3445
QY	3686	gtcacgactcagacccgcgtactctgaccctcagataaaggctcgaagtgcgcattgtgt	3745
Db	3446	GCCATGACACTGACCCCGCTACGCCGAGGCCACTGTGAAGGCTCGTGTGGCCGAGCTGT	3505
QY	3746	atctacctctgattggtattatcatctgnaaacctgtfacctcagctgttatgtatttcagaaa	3805
Db	3506	ACCTGCCACTGCTTTCGATTGCAACGGGATACCTTGCCACGGCTGCATGACTTTGCTGAGG	3565
QY	3806	ctcacaaatcaacgaggaagac-----caatttgatagccactgatgatataaagtg	3859
Db	3566	GCCCCAGGTCAAGCGGTCAAGACTGGCCCTCAATGCTTGACTCAGACACAGAAAGGCCGAAGGGG	3625
QY	3860	agagcggaaagtaagtaagccagagaccgttgcacatgycacatgcagagacatcogtccctc	3919
Db	3626	ACATTGGCGGTACCATCAACCCCTCTGTGGCCATGGCCATTGTGTTGGGCCCTTAGCCC	3685
QY	3920	aactacaagaagccttgcagtttccctccctcagtcacaacgagtygcagcgaacacactacct	3979
Db	3686	CTGGCTCCCGGGCCAGCATCTCCCAAGGGGCCAACCAACGGCTTCTGCGCAGAGCTGTGCC	3745
QY	3980	ttcagcagaatacgaagtcgaagccctttagatctgctactttaggttctcaaaaatgcag	4039
Db	3746	TCTGTGTGATGATCAAGCCGACCTTGTGCGGTGTGTCTGTGGTGTGTAATAAACACCG	3805
QY	4040	atgaaacagttctacagaagtggtttacagatctctcagtcctcagtcctgcagcttaaacccgctat	4099
Db	3806	AGCCGGCGCTCCTGCAGCGCTGGGCCACTGACACTCCCCCAAGCTGGGAGCGTCTGT	3865
QY	4100	tagatctgcttatctctgtgtgtcttgcctttagtataaagggaaaaaagtgtttgaac	4159
Db	3866	TGGACTTGTCTGTACCTTTGGCCTGGCTGCCTTTGAGTACAAAGGGGAAAAAGCCCTTTGAAC	3925
QY	4160	gaatgaatagcttgaccctttaagaatacaaaagacatgagagcaaaagcttgaagaagcta	4219
Db	3926	GCATCAACAGCCTCACATTTCAAAAAATCTCTTGATATGAAGGCGCGGCTAGAGGAAGCCA	3985

QY	4220	tctcttgagagcatatggtgcccagggcaagaatatgtatcagcgcggaagccgagacagcttcgaga	4279
Db	3986	TTCTGGGTACCATCGGAGCTCGACAAGAAATGTTTCGGCGGAAGTCGTGAGAGG-----	4038
QY	4280	gaagcccatctgtgaagtgcccttctggaagtcagaaaaaattgaggtgaggaatatga	4339
Db	4039	-----AGCCCGTTTGGGAATCCAGAGAACGTGCGCTGGCGGAAGACGCTCA	4084
QY	4340	ctcactggcgtaaaaaacacagagaagaagcttgacaatatcaagagcagagattgaaccgaag	4399
Db	4085	CACACTGGAGCAAAACCTCAGACCCGCGTGGACAGAACCAAGGATGAATGGAACACGAGG	4144
QY	4400	cactgatctgatgtgaacctgtgctacagaaagcaacctaatcatctttagatatcatagaga	4459
Db	4145	CCTTGGTGGGAAGGGAACCTGGCAACCGAGGCAAGCCTAGTGGTTCGTGCACACTGGAGA	4204
QY	4460	ttgtgtgtcagaccggttctgttaacggaatccaaaagagagcatctctgtgtgagtctaa	4519
Db	4205	TCATCGTGGCAGACGGGTGATGCTTTCAGAGACCCCGGAGAGACGGCTTTGGGGGCAGTCTGA	4264
QY	4520	aagtgctactacacagcatgtgctgttaaccaaaagtgacgttatactacaacatgttttg	4579
Db	4265	AGGTTGTGCTGTACAGCCTGGGCAGTGGCCACAGAGTGCCCTCTTCTTGACAGCATGGCCCTGG	4324
QY	4580	ctacacagagaagccctgtgttccaagtttcctgaactcttatttgaagaagagacagagc	4639
Db	4325	CCACCAGAGGGGCCCTGTGTCCAAGTTCGCCGAGCTGCTGTTGAGGAGGACACGGAGC	4384
QY	4640	agtggtctgatttatgtcctcaggtcttcgcacactgtagcagtagacatcggtacaaatc	4699
Db	4385	TGTGTGCCGACCTGTGCGCTGAGGCTCCTACGACACTGTGGCAGCCGCAFCAGCACCATCC	4444
QY	4700	ggtcacacccccagtgccctcccttaccctactaatgagygcaaaactttagattggaata	4759
Db	4445	GCACGCACGCCACGCGCTCGCTGTACTGTCTCATGCGACAGAACTTCGAGATCGGCCACA	4504
QY	4760	acttgtccaggtttaaattgcaggtlaccaatgttcactatccctcctgtgtggcacatctc	4819
Db	4505	ACTTTGCCCGGTGAGATGTCAGGTACCATGTCTCTCTCGTCCCTGGTGGGACGACGCGC	4564
QY	4820	agaattttaatgaagaattccttaaacggttctcctaagaactatatgtacatatgctgaag	4879
Db	4565	AGAACTTCAGTGAAGAGCACCTGCGACGTTCACTCAAAAACCATCTCACTATATGCTGAGG	4624
QY	4880	aagactctggaattgagygaaacaacattctcgtatcaggtccagatctgtgtttcaatc	4939
Db	4625	AGGACATGGGGCTGCGGGACAGCACTTCGACAGCAGGTCCAGGACTGATGTTCAACC	4684
QY	4940	tccatatgatcttcttgatactgtgaaaaatgaagaacaccagagagatcctgaatgt	4999
Db	4685	TGCACATGATCCTGACGCGACACGGTGAAGATGAAGGAACACACAGAGAGCCCTGAGATGC	4744
QY	5000	tgattgatctaattgtacagaattgtccaaggttaccagacctctccaga---gcatatga	5056
Db	4745	TCATCGACCTCATGTACAGAAATGCCCCGGGGCTATCCAGGGGCTCACCGGACCTTCGGCTGA	4804
QY	5057	cctggttgcaagaacatgycagygcaagcactcagaacgaagcaatcatatgtgaagctgcac	5116
Db	4805	CCTGGTTGCAGAACATGGCCGGGAAGCACCGCGGAGCTGGGCAACCAACGCGGAGGCCGCC	4864
QY	5117	agtgcttagtccactcagcagc 5138	
Db	4865	AGTGCATGTTGCACGGCGCGC 4886	

RESULT	13				
AK024436					
LOCUS	AK024436	4577 bp	mRNA	linear	PRI 29-SEP-2000
DEFINITION	Homo sapiens mRNA for FLJ000026 protein, partial cds.				
ACCESSION	AK024436				
VERSION	AK024436.1	GI:10440380			
KEYWORDS	fls (full insert sequence).				
SOURCE	Homo sapiens adult spleen cDNA to mRNA, clone:as00026.				

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 4577)
AUTHORS Ohara,O., Nagase,T., Kikuno,R. and Okumura,K.
TITLE The nucleotide sequence of a long cDNA clone isolated from human spleen
JOURNAL Published Only in DataBase (2000) In press
REFERENCE 2 (bases 1 to 4577)
AUTHORS Ohara,O., Nagase,T., Kikuno,R. and Okumura,K.
TITLE Direct Submission
JOURNAL Submitted (24-AUG-2000) Osamu Ohara, Kazusa DNA Research Institute, Department of Human Gene Research, 1532-3, Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:cdnainfo@kazusa.or.jp, URL:http://www.kazusa.or.jp/NEDO, Tel:81-438-52-3913, Fax:81-438-52-3914)
COMMENT NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction, 5' - & 3'-end one pass sequencing and clone selection: Kazusa DNA Research Institute.
FEATURES
source
1..4577
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="as00026"
/tissue_type="spleen"
/dev_stage="adult"
/note="vector:pBluescriptII SK plus"
1..3544
/gene="FLJ00026"
<1..3544
/gene="FLJ00026"
/note="Start codon is not identified."
/codon_start=2
/product="FLJ00026 protein"
/protein_id="BAB15726.1"
/db_xref="GI:10440381"
/translation="KNIMSSKIADRNCSRMSYCYSGSSDAPSPAPRPASKKHFEH
LAIQMVSTGMVRETYEKYAMKFFELLVKSMAQHVHNMCKRDSFRTRFSDRMDIT
TIVNVTSSEIALLVKPOKENEQAEKMNISLAFLYDLISMDRGVENLIRHYCSOL
SAKLSNLPILISMRLFLRLCSHEHYLNLFPMNADTAPTSPSPSISSQNSSSCSS
FODOKIASMFEDLTSEYRQOHFLTGLFTTELAALDAEGEGISKYORAVSAIHSLS
HDLDPRCVKPEYKVKIAALYPLVGIILDALPOLCDEFTVADTRRYRTSGSDEOEGAG
AINONVALAIGNNENLKTSGVLSLPYKQYNMLNADTNLMLICELMIMKNADQSL
IRKWIADLPSTQLNRILDLFLICYLCEFYKGKQSSDKVSTQVLQKSRDYKARLEAL
RGEARGEMMRRAFGNDRFGLNENLRMKKEQTHWRQANEKLDKRAELDQALISG
NLATEAHLILDMQENIIQASSALDCKSLGGVLRVLVNSLNCQSTYLYTHCFATL
RALIAKFGDLFEFEVQCFDLCQVLRHSCSSMDVTRSQCATLYILMRSFGATSN
FARVKQVMTSLASLVGRAPDFNEHLRSLRTILAYSEEDTAMQMPPTQVEELIC
NLNSILYTVKMRFEQDEPEMLMDLMRIAKSYQASPDRLRTLQNMAREKHTKKCYT
EAAMCLVHAALVAEYLSMLEDSYLPVGSYSFQNISSNVIEESVSEDTLSPDEGV
CAGQFTESGLVGLLEQAAELFSTGGLYETVNEVYKLVIPILAHREFRKLTLTSHKL
QRAFDSIVNKDHRMEGTLYERVGFGSKFGDLDEQEFYKEPATIKLPEISHRLAFY
GQCFGAEEVEVITDSTPVDKTKLDPNKAYIQITVEPYFDEYEMKDRVYFEKNFNL
RFMYTPTFTLEGRGELHEQYRNTVLTTHMAFPYIKTRISVIQEEFVLTPIEVAI
EDMKKKTLLQAVAINQEPDPDAKMLQMVNQSVGATVNOGPLEVAQVFLAEIPADPKLY
RHHNKLRCFKEFTMRGCEAVEKRNRLITADQREYQOELKKNNYKLENLRPMIERKI
PELYKPIFRVESQKRDSFHRSSFRKCEQLSQGS"
BASE COUNT 1297 a 1037 c 1077 g 1166 t
ORIGIN
Query Match 17.9%; Score 1223.6; DB 9; Length 4577;
Best Local Similarity 62.3%; Pred. No. 4.4e-275;
Matches 2122; Conservative 0; Mismatches 1219; Indels 67; Gaps 10;

Db 166 GGTGAGAGAAACAGTCTTCAAGTATAGCCTGCTTCTTTGAGCTTCTGTTGAAAAGCAT 225
QY 3078 ggtgaccattataactttaatgataaacttgagctccaaggaagaagtcgtttccaga 3137
Db 226 GGCCGACGACGTACATACATGACACAACGGGACAGTTTTCGAGGACTCGTTTCTGTA 285
QY 3138 acgttcattgatgacattgcagctctgttcagcaacgattgtcagatatagttcacg 3197
Db 286 CCGTTTCATGATGACATATACTACTATGTTAATGTGGTCACTCCGGAATTCAGCCCT 345
QY 3198 att-----tcagaagcacacagaaatggttagagacatcaatacaagccttgcat 3248
Db 346 TTTAGTAAACCACAGAGAAGAAATGACAGCGGGAAGATGACATCAGCCTGGCTTT 405
QY 3249 cttctcaatgatctgtgtctgttatgagacagagattgttttagcctataaagtc 3308
Db 406 CTTCTTGATGACCTTCTCTCCCTCATGGATCGGGGCTTTGTTTAACCTCATCAGACA 465
QY 3309 ctgctataacagtgcttccaagcttactcaatcaccgaatccagtgctgtgtc 3368
Db 466 TTATTGACGCCAGCTGTCAGCCAAGCTCAGTAACCTTCCAACGGCTCATT-----TC 516
QY 3369 ctgagcctgatttctcagaatcatctgcagtcatgagcattatgtacattaaact 3428
Db 517 CATGAGGCTAGAGTTCCTGAGAAATCCCTGTAGCCATGAGCATATACCTCATGAACT 576
QY 3429 accctgcagcttacttaccacctgcacatccacaccttcgtttctctcgaacatc 3488
Db 577 -----TTTTTTATGAATGCTGATACTGCTCCACATCTCCTGTCTTCATATCTTC 630
QY 3489 tcagagttctgatttctcagaatgtacaagaaccaagaatccaatatgtttgaatt 3548
Db 631 CCAGAACTCAAGCTCTCTGCTCCAGCTTCCAGGACCAGAAAGATCGCCACATGTTGATCT 690
QY 3549 atccgtgccttcgcgccaacagcattatttggcaggacttgtttaacagagctgctgt 3608
Db 691 GACTTCCGAGTACCGCCAGCAGCAGCAGCTCTCACCAGGCTCCTCTTCAAGAACTGGCTGC 750
QY 3609 catttagaccctgatgctgaagagctgtttgattgcatagaagaagtcacaatatgt 3668
Db 751 TGCCCTGATGCCGAAGGGGAAGGAATCAGCAAAAGTACAAAGGAAGCTGTCAGTGAAT 810
QY 3669 acacaatttactctccagtcacgactcagaccgcggtactctgcacccctcagataaagtc 3728
Db 811 TCACAGCCTGCTAAGTTCTTCACGACCTGGACCCACGCTGTGTCAACCAAGAGTGAAAGT 870
QY 3729 tcgagtgccatgttgttatctacacctgattggtatatacatggaactgtlactcagct 3788
Db 871 CAAAATCGCCGCCCTTTTAACTTACCTTTAGTTGGCATCATTTTGGATGCTTTGGCACAGCT 930
QY 3789 gtatgatttacagaagaactcaacaatcaagaagaagaaccaattgtatagccaatgatga 3848
Db 931 CTGTGACTTTTACAGTTTGACAGATCTCGAGATACCGCACACAGTGG-----CTCGGA 981
QY 3849 ttatgaagtgaagcgggaagtatgataagccagaccgttgccatggcaatcgcagggac 3908
Db 982 TGAAGAACAAGAGGAGCCGGTGCCATTTAACCAAGATGTGGCTTGCCCATAGCAGGAA 1041
QY 3909 atcggtccctcaactaacaagcgctgcaagtttcctcctcaagtcgaacgagtgcgagca 3968
Db 1042 TAATTTCAATTTGAACAACAAGT-----GGAAATAGTGCTGTCTTCCTTGCCCTATAAGCA 1095
QY 3969 acacactacccttccagcagaatcaagaatcgaaagccttttgatcgtctactttggtctc 4028
Db 1096 GTACAACATGCTGAACGGGACACACTACTTCGCAACCTCATGATCTGCTTCCCTGTGATCAT 1155
QY 4029 caaaaatgcagatgaacaagctctacagaagtgtttacagatctctcagctctgcagct 4088
Db 1156 GAAAAATGCTGATCAGAGCCTCATTTAGGAAGTGGATTGCTGACCTGCCATCAACGACAGCT 1215
QY 4089 aaacggtattagatcgtcttatctcgtgtgtcgttgcttgagtlaglataaaggaaanaa 4148
Db 1216 CAACAGGATTTTGAATCTACTTTTCATCTGTGTGTATGTTTGAATATAAGGAAACA 1275

QY	4149	agtgtttgaacgaaatgaataagcttgacccttlaagaaatcaaaaacatgagcgaaaagct	4208
Db	1276	GAGTTCGTACAAAGTCACTAACCCAAAGTCCCTGCAGAGACTCAAGGATGTCAAGGCCGGCT	1335
QY	4209	tgaagaagcattctctgggagcataggtgccaggcagaagaatgytacgcgaagccgag	4268
Db	1336	GGAAGAGGCTTTGCTCCCGTGGGGAAGGGGCCAGAGGGAGATGATGCGCCGCCGGCTCC	1395
QY	4269	acagctcgagagaagcccactctggaagtgcctttggaagtcaagaaattgaggtgag	4328
Db	1396	AGGGAACGACCATTTCACAGGCCCTAA-----ATGAAAATTGTAGATGGAA	1440
QY	4329	gaaagatatgacctcaactygcgtcaaaaaacacagagaagcttgcacaatcaagacagagat	4388
Db	1441	GAAAGACGACACACTTTGGCGGCAAGCTAATGAGAAGCTAGATAAAACAAGGCCGAGTT	1500
QY	4389	tgaacacgaagcacctgatgtatgynaacctggtctacagaagcnaacctaatcttaga	4448
Db	1501	AGATCAAGAAGCCTTGATCATAGTGCCAATCTGGCTACAGAAGCACATTTAATCATCTTGA	1560
QY	4449	tacatlagagattgttgttcagaccglttctgtbaacygaatccaagagagcatcttg	4508
Db	1561	TATGCAGGAAACATTTTCCAGGGCAGCTCGGCTGTGACTGTAAGACACAGCCTTCTGGG	1620
QY	4509	tggagtgctaaaagtgctactactacacagcatgycgtctgtaaccaaaagtgcagttatcta	4568
Db	1621	AGGTGTTCTGAGGGTGCTGTGTGAATTTCTCTGAACGTGTATCAGAGTACCACCTACTGAC	1680
QY	4569	acactgtttgtctacacagagagcccttggttccaagtltcctgaactctlatttgaaga	4628
Db	1681	TCACGTCTTGCAACACTCCGTGCTCTCATCGCCAAGTTGGAGACTTACTCTTGAAGA	1740
QY	4629	agaacagagcagtgctgatttatgctcaggtcttcogacactgtagcaqtagcat	4688
Db	1741	GGAGGTGGAACAGTGTTCGACCTATGTACACCAGTCCCTGCACCACTGCAGCAGCAGCAT	1800
QY	4689	cggtacacatacgttcacacccccagtgctccctttaactactaatgagcgaaaactlta	4748
Db	1801	GGATGTCACCCCGGAGCCCAAGCCTGTGCCACCCCTTAACTCCTCATGAGGTTCAAGTTTTGG	1860
QY	4749	gattgggaataacttgcaggggttaaagtccaggtaccaatgtcatactatcctccttgt	4808
Db	1861	AGCCACCAGTAATTTGCCAAGAGTAAGATGCAAGTAAACATGTCCCTGGCATCTTTGGT	1920
QY	4809	gggcacatctcagaattttaatgaagaattcttaagacgttctctaagacataattgac	4868
Db	1921	GGGAAGAGCACCAGACTTTAATGAAGAGCACCTGGAAGATCCTTGAGGACAAATTTGGC	1980
QY	4869	atatgtgaagaagatcttggaaattgagggaaacaatttcctgtatcaggtlccagatct	4928
Db	1981	CTATTCAGAAGAGGACACACAGCCATGCAGATGACTCTTTCCACCACCGGTGAGGAAGT	2040
QY	4929	ggtttccaatcctccatatgattcttctgatalactgtgaaaaatgaaagaaacaccaggagga	4988
Db	2041	TCTCTGTAATCTGAATAGCATCTTATATGACACAGTGAATAATGAGGGAATTCAGGAAGA	2100
QY	4989	tcctgaaatgttgattgatactaaatgtlacagaattgccaaaggtttaccagacctctcaga	5048
Db	2101	TCCTGAGATGCTTATGATCTCATATGTACAGAAATGCCAAGAGTTACCAGGCAATCTCTGA	2160
QY	5049	--gcattgacctggttgcagaacatgycagcgcaagcactcagaacgaagcaatcatgc	5105
Db	2161	TCTGCGGTGACCTGGCTCCAGAACATGGCAGAGAAACACACCAAGAAGAGTGTCTACAC	2220
QY	5106	tgaagctgcacagtgctctagtcacctcagcagcaactgttgtctgaatatattgacatgct	5165
Db	2221	GGAGGCTGCCATGTGCTGTGTGCACGCCGCTGCGTTAGTGCGTAGTATCTGAGCATGCT	2280
QY	5166	ggaggaaccggaatatattctcctgtgtggatgtgtaaacattcagaatatattcatctaatgt	5225
Db	2281	GGAGGACCACAGCTACTGCCCCGTGGCGAGTGTACGCTTCAGAAATATTTTCCAAATGT	2340

QY	5226	tttagaagaatctgcggtctcagatgatgtgtatctccagatgaaagatctgtc	5285
Db	2341	GCTGGAGAGTCTGTGCTCTTGAGGACACCCTGTCACTGACGAGGATGGGTGTGGC	2400
QY	5286	tggaanaatacttactgaagtcagagactgtggattactggaacaagcagctgtccct	5345
Db	2401	AGGCCAGTACTTCACCAGAGTGGCCTGGTAGGCCCTCTGGAGCAGGCCGGAGCTCTT	2460
QY	5346	ctctatgctgycatgtatgaagcagttaatgaagtttacaagttacttatccattca	5405
Db	2461	CAGCAGGGAGGCTTATATGAGACAGTTAATGAGGCTTACAAGCTGGTCATCCCATCT	2520
QY	5406	tgaagctaatcgggatgcaagaactatccacaattcatatgttaacttcaagaagcatt	5465
Db	2521	AGAAGCCCATCGAGAAATCCGGAAGCTGACACTCACTCACAGCAAGCTGCAGAGACCTT	2580
QY	5466	cagcaaaattgttcatcagaagtcgtctgggagcgygatgtttgycacctatcttctgt	5525
Db	2581	CGACAGCATCGTTAACAAAGATCAT-----AAGAAATGTTTGAACCTACTTCCGAGT	2634
QY	5526	tgtttttaaagaaacaaagttcgggatttgatgaaacaagaattgtttaaaggagcc	5585
Db	2635	TGGTTCTTTGGATCCAAATTTGGGATTTGGATGAACAGGAATTTGCTACAAAGAGCC	2694
QY	5586	tgcataccaacaaacttgcaagatatccacagattgagggatttcaagaaagaatt	5645
Db	2695	TGCAATTACCAAGCTTCTGAGATCTCACATAGACTAGAGCATTTTATGGTCAATGTTT	2754
QY	5646	tggagaggatgtgttgaagtaatcaagaactctaactcgttagacaagtgtaattaga	5705
Db	2755	TGGTGCAAAATTTGTGGAAGTGATTAAGACTCCACTCCTGTGGACAAACCAAGTTGGA	2814
QY	5706	tcctaacaaggcatatatcagattacctatgttggagccataacttgacacatatgagat	5765
Db	2815	TTCTTAACAAGGCTTACATACAGATCACTTTTGTGGAGCCCTACTTTGATGATATGAGAT	2874
QY	5766	gaagagcagaatcaacctalttcgacaaaaattacaactctgcgcattcatatgttctac	5825
Db	2875	GAAAGACAGGGTTCACATACTTTGAGAAGAAATTTCAACCTCCGAGGTTCATGTACACCAC	2934
QY	5826	acccttactttagatgycgctgcccattggygaactlcatgaaacattcaaaagyaagac	5885
Db	2935	CCCGTTCACCCCTGGAGGGCGGCTCGGGGAGAGCTGCATGAGCAGTACAGAGGAACAC	2994
QY	5886	catctgactaagctcatatgcttcccttatataaaacaaggtcaatgtcactcataa	5945
Db	2995	AGTCTGACCACTATGCAAGCGCTTCCCTTACATCAAGACCAGGATCAGCGTCATCCAGAA	3054
QY	5946	agaagagatcatcttaacaccaattgaagttgtctattgagacatgcagaaaaagacaca	6005
Db	3055	GGAGAGATTGTTTGTGACACCGATTTGAGTTGCCATTGAAGACATGAAGAGAAGAACCT	3114
QY	6006	ggagttggcatttgcaacacatcagaatcccgagaccccaaaatgcttcagatgttact	6065
Db	3115	GCAGTTAGCAGTTGCCATTAAACCAGGACCCGCCCTGATGCAAAAGATGCTTCAGATGGTCT	3174
QY	6066	ccaaggaatctttagcaccacacagtgaaatcaggggccttgygaagttgccaggttct	6125
Db	3175	GCAAGCTCTGTGGAGCTTACTGTAAATCAGGGACCACTGGAAGTAGCCCAAGTGTTTYT	3234
QY	6126	gtctgaataatcagtagtgaccacaaagctctcagacatcataataaactgcagcttgc	6185
Db	3235	GGCTGAATTCCTGCTGATCCAAAACCTTATCGACATCACAAACAAGTTGAGGTTATGCTT	3294
QY	6186	taagattttactaaaaggtgtgaagatgyccttaagaanaaataagagcttaatttggcc	6245
Db	3295	TAAAGAAATTCATCATGAGATGTGTGAAGCTGTAGAGAAAAACAAGCGTCTCATCACGGC	3354
QY	6246	ggttcaaaaagagtatcaaaaggaatttgggga--aactatcttgccttaagagggcct	6303
Db	3355	AGACCAGAGGGGAATATCAGCAGGAACCTCAAAAAGAACTATAACAAGCTAAAGAAGAACCT	3414
QY	6304	acaagccctagatc--acagaaagtcctcagttalccaagccagattt	6349


```

|||||
Db    3415 CAGGCCAATGATCGAGCGGAAATTCAGAACTGTACAGCCCAATATT 3462
|||||
RESULT 14
BC008335
LOCUS      2825 bp      mRNA      linear      PRI 12-JUL-2001
DEFINITION Homo sapiens, clone IMAGE:3506202, mRNA, partial cds.
ACCESSION BC008335
VERSION    BC008335.1   GI:14249910
KEYWORDS
SOURCE
ORGANISM   human.
            Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 2825)
            Strausberg, R.
            Direct Submission
            Submitted (25-MAY-2001) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
REMARK
COMMENT    NIH-MGC Project URL: http://mgc.nci.nih.gov
            Contact: MGC help desk
            Email: cgapbs-rt@mail.nih.gov
            Tissue Procurement: DCTD/DTP
            cDNA Library Preparation: Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
            DNA Sequencing by: Genome Sequence Centre,
            BC Cancer Agency, Vancouver, BC, Canada
            info@bcgsc.bc.ca
            Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
            Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
            Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
            Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
            Ness, Pawan Pandoh, Anna-Liisa Priabhu, Parvaneh Saeedi, Jacqueline
            Schein, Duane Smalhus, Michael Smith, Lorraine Spence, Jeff Stott,
            Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
            George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: http://limage.lnl.gov
Series: IRAL Plate: 23 Row: d Column: 12.
Location/Qualifiers
1. 2825
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3506202"
/lissue_type="Kidney, renal cell adenocarcinoma"
/clone_1ib="NIH_MGC_14"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
<1..2585
/codon_start=3
/product="Unknown (protein for IMAGE:3506202)"
/protein_id="AAH08335.1"
/db_xref="GI:14249911"
/translation="QRSLASMLDSDTEGEDDIAGTINPSYAMAIAAGPLAPGSRASI
SQGPTASRAGCALSAESSRTLACVAVLKNTEPALLOQWATDLTPOLGRLLDLTY
LCIAAFEYKGRKKAFFERINSLTFKKSIDMKARLEEAIGTIGARQEMVRSRERSPEGN
PENVRWRKSVTHWKOTSDRVDTKDEMEHEALVEGNLATEASLVLDLEIIVQTVML
SEARESVLGAVLKVLYSLGSAOSALFLOGLATORALVSKEPDLFEEDTELCAIDC
LRLRHGGSRISTIRTHASASLYLMRONFEIGHNFARVKMOYTMSSLSLVGTQONS
EHLRRLSKTLITYAEDMGIRDSTPEAQVODIMFNLMHTLDTYKMKHEQEDPEMLI
DLMYRIARGYOGSPDLRLTWLQNMAGKHAELGNHAEAAQCWYHAALVAEYLALLEH
RHLPGCVSFQNISSNVLLEESASIDILSPDEEGFCSGKHFTBLGLVGLLEQAAGYFT
MGGLEYAVNEVYKNIPLLEAHRDKKLAAVHGKIQEAFKTMHQSSGWERVFGTYFR
VGFYGAHFGLDDEQEEFVKERSITKLAIEISHRLEEFYTERFGDYVEIIKDSNPYDKS
KLDSQKAYIQITTYVEPYFDITYELKDRVTVFDRNYGLRTFLCTPPTPDGRAHGELPEQ
HKRKTLSSTHAFPYIKTRIRVCHRETVTLTPEVAIEDMOKKTRELAFATEQDDPPDA
KMLQWLVLOGSVGPTVNOGFLVAQYFLAIEIPEDPKLFRHHNKRILCFKDFCKCEDAL
RKNKALIGPDQKEYHRELERNYCRLEALQPLLTQRLPOLMAPTPPGLRNSLNRASFR
KADL"
```

```

BASE COUNT      641 a      835 c      827 g      522 t
ORIGIN
Query Match      16.4%; Score 1119.4; DB 9; Length 2825;
Best Local Similarity 66.0%; Pred. No. 1e-250;
Matches 1694; Conservative 0; Mismatches 841; Indels 30; Gaps 4;

QY    3853 gaaagtgaagagcggaagtatgataagccagaaccgttgcgatgcaatgcagygacatcg 3912
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db     48  GAAGGGACATATGCGGGTACCATCAACCCCTCTGTGGCCATGGCCATTGCTGTGGCCCC 107

QY    3913 gtccctcaactaacaagcgcttgcagtttctcctcactacgtcaacgaagtgcagccaac 3972
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db     108 CTAGCCCCCTGGCTTCCCGGCCAGCATCTCCAGGGGCCACCACCAAGGCTTCTCGCCAGGC 167

QY    3973 actaccttcaagcagaatcaagtcgaagccttltgatctgtctacttlttggttctcaa 4032
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db     168 TGTGCCCTCTCTGCTGAGTCAAGCCGGAACCTTGCTGGCGTGTGTGTGGTGCTGAAA 227

QY    4033 aatgcagatgaacacagttctacagaagtggttacagatctctcagtcctgcagctaac 4092
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db     228 AACACCGAGCGCGGCTCCTGACAGCGGTGGGCCACTGACCTGACACTCCCCACAGTGGGA 287

QY    4093 cggctattagatctgtcttatctctgtgtgtcttgcctttagtataaaggaaaaaagt 4152
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db     288 CGTCTGTGTGACTTGTCTGTACTCTTGCCCTGGCTGCCCTTGTGAGTACAAAGGGAAGGCC 347

QY    4153 ttgaacgaatgaatagcttgcctttaaagaatcaaaaagacatgagcaaacgttgaa 4212
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db     348 TTGAACGCATCAACACAGCCTCACATTCAAAAATCTCTGGATATGAAGCGCGGCTAGAG 407

QY    4213 gaagctattcttgggagcatagtgccaaggaagaatggtcagcggaagcgagagaa 4272
      ||||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db     408 GAAGCCATTCTGGGTACCATCGAGGCTGCACAAGAAATGCTTGGCGGAAGTCTGAGAGG 467

QY    4273 ctcgagagaagcccatctgtgaagtccttggaaagtcgaagaaattgagtgagga 4332
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db     468 -----AGCCCGTTTGGGAATCCAGAGACGTGCGCTGGCGGAAG 506

QY    4333 gatatgaactactgcygtcaaaaacagagaagccttgacaaatcaagagcagagatgaa 4392
      || ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     507 AGGTCACACACTGGAAGCAAACTCAGACCGCGTGGACAAAGACCAAGATGAATGGA 566

QY    4393 cacgaagcactgattgatgaaacctggtctacagaagcaacctaatcatltagataca 4452
      ||||| || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db     567 CACGAGGCCCTTGGTGAAGGGAACCTGGCAACCGAGCAAGCCTAGTGTCTGACACACA 626

QY    4453 ttagagattgtgttcagaccgttctgttaacggaatccaaagagagcatctctgtgga 4512
      || ||||| || ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db     627 CTGGAGATCATCGTGCAGACCGGTGATGCTTTCAGAAAGCCCCGGGAGAGCGCTTGGGGCA 686

QY    4513 gtgctaaaagtgtactacacagcattggttcaaacagttcgtcaactcttattctacaac 4572
      ||||| ||| ||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db     687 GTGCTGAAGGTTGTGCTGTACAGCCTGGGGCAGTGCCACAGATGCCCCCTCTTCTTGACAGAT 746

QY    4573 tgtttgtacaacagagagccttggtttcaaaagttccttgaactcttatttgaagaagag 4632
      || ||| ||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db     747 GGCCGTGGCCACCCAGAGGGCCCTGTGTCCAAGTTCCCGGACCTGCTGTTGAGAGGAGAC 806

QY    4633 acagagcagtgctgatttattgctcagagcttctcagacaactgttagcagtagcatcgt 4692
      || ||||| ||||| ||| ||||| ||||| ||| ||||| ||||| ||||| ||||| |||||
Db     807 ACGAGCTGTGTGCCGACCTGTGCTGAGGCTCTACGACACTGTGGCAGCCGCAATCAGC 866

QY    4693 acaatacgtcacacccccagctgctcccttaactactaattgagggcaaacatttgagatt 4752
      || ||| ||| ||| ||| ||||| ||| ||||| || ||| || ||| ||| ||||| |||||
Db     867 ACCATCCGCAACGACGCGCAGCGCCTGCTGTACTCTGCTCATGCGGACAGAACTTCGAGATC 926

QY    4753 gggataaacttgcagggttaaatgcaggttaccaatgtcactatcctccttgtgtggc 4812
      || ||| ||||| ||| ||| ||||| ||| ||||| || ||| || ||| ||| ||||| |||||
Db     927 GGCCACAACTTTGCCCGCTGTGAAGATGCAGGTCAACCATGTCTCTCGTCCCTGTGTGGGG 986

QY    4813 acatctcagaattttaatgaagaattcttaagaagcttctctaaagactatattgacatat 4872
```

Db	987	ACGACGACGAACTTCAGTGAAGAGCACCTGCGAGCTTCACTCAAAACCATCCTCACCTAT	1046
QY	4873	gctgaagaagatctggaaattgaggaacaacatttcctgatcaggtccagatctgtt	4932
Db	1047	GCTGAGGAGACATGGGGCTGCGGACAGCACTTCGCAGAGCAGGTCCAGGACCTGATG	1106
QY	4933	ttcaatctccatatgatcttctctgatactgtgaaatgaagaaaccaggagatcct	4992
Db	1107	TTCAACCTGCACATGATCCTGACGAGACACGGTGAAGATGAAGAACACAGAGACCCCT	1166
QY	4993	gaatgttgattgtatctaattgtacagaattgccaaggtttaccagacctctccaga--g	5049
Db	1167	GAGATGCTCATCGAAGCTCATGTACAGAAATGGCCCCGGGCTACCAGGGCTCACCGACCTT	1226
QY	5050	cgattgacctgtgtcagaacatgycagggcagcaactcagaacgaaagcaatcatgtgaa	5109
Db	1227	CGGCTGACCTGGTTGCAGAACATGGCCGGGAAGCACGCGGAGCTGGGCAACCAAGCCGAG	1286
QY	5110	gctgcacagtgctcagtcacactcagcagcaactgtgtcgaatatattgagcatgtgag	5169
Db	1287	GCCGCCAGTGCATGGGTGCACGCGCGCCGCTGTGGCTGAGTACCTGCGCCTGCTCGAG	1346
QY	5170	gaccggaatatcttcctgttggatgtgttaacattcagaatatattcatctaatttta	5229
Db	1347	GACCAACCGCCACCTGCGCCGTGGGCTGGCTTTCCTTCCAGAACATCTCATCCAACGTGCTA	1406
QY	5230	gaagaatctgcggttcagatgtatgtgtatctccagatgaagaaggtatctgtcttga	5289
Db	1407	GAGGAGTCCGCCATCTCCGACGACATCCTGTGCCCCGACGAGAGGGCTTCTGCTCCGGG	1466
QY	5290	aaatacttactgagtcaggactgttggatattctggaacaagcagctgtcctctct	5349
Db	1467	AAGCACTTCACGTGAGTGGGGCTGTGAGGGTGTCTGGAACAGGCAGCCGGCTACTTCACC	1526
QY	5350	atgctgycatgtaataagcagtttaataagtttaacaaagtacttattcctattcatgaa	5409
Db	1527	ATGGGGGGGCTTACGAGGGCGGTGAATGAGGTCACAGAACCCTCATCCCATCTCGGAA	1586
QY	5410	gctaatacgggatgcaaaagaaactataccacaattcatgttaacttcaagaagcatcagc	5469
Db	1587	GCCCCACCGTGACTACAGAAGAGCTGGCCGCGGTGCACGGCAAACTGCAGGAGGGCCTTCACC	1646
QY	5470	aaaattgttcacagagtaactgctgtyggaagcgatgttggcaacctatttcgttgtgt	5529
Db	1647	AAGATCATGCACACAGAGTTCGGGCTGGAGCGCGTGTTCGGAGCGTATTTCGCGTGGGC	1706
QY	5530	tttataagaaacgaagtcggygatttggatgaacaagaattgtttacaagagcgctgca	5589
Db	1707	TTCTACGGCGCCACTTCGGTGACCTGGATGAGCAGAGTTTGTGTACAAAGAGGCCATCG	1766
QY	5590	ataaccaacttgcagagatatctcacagatgaggagatttaagagaaagatttga	5649
Db	1767	ATCACGAAGCTGGCAGAGATCTCACACCGGCGGAGAGATTCTACACGGAGAGATTGGC	1826
QY	5650	gaggaatgtgtgaagtaatacaagactctaactcctgttagacaagtgtaaatagatcct	5709
Db	1827	GACGACGTCGTTGAGATTATCAAAAGACTCTAACCCCTGTGGACAAGTCCAAGCTTGACTCA	1886
QY	5710	aacaaggcatatatcagaattacctatgtggaagccatacttggacacatatgagatgaag	5769
Db	1887	CAAAAGGCTTACATCAGATCAGTATGTGAACCGTACTTTGATACTACGAGCTCAAG	1946
QY	5770	gacagaatcacctatttcgacaaaattacaactctcgtcgattcatgtactgtacaccc	5829
Db	1947	GACCGGGTGACCTACTTTCACCGCAACTATGGGCTTCGCACATTCCTCTTCTGCACGCCG	2006
QY	5830	tttactttagatggcggtgccatggygaacttcatgaaacaattcaaaaggaagaccatt	5889
Db	2007	TTACAGCCCGGATGGCGCGCACACGCGGAGCTGCCCGAGCAACACAGCGTAAGACGCTG	2066
QY	5890	ctgactacgtctcatgtccttcttataataaaacaaggytcaatgttcactcataagaa	5949

Db	2067	CTCAGCACCGACACGCGCTTCCCTACATCAAGACTCGCATCCGTGTGTGCCACCGGGAG	2126
QY	5950	gagatcatcttaacaccaattgaagttgtctatgtagacatgcagaaanaagacagagag	6009
Db	2127	GAGACGGTGTCTGACCGCCAGTGGAGGTGGCCATTCAGACATGCAGAGAAGACACGGGAG	2186
QY	6010	ttggcatttgcacacatcagatcccgagacccccaaaatgcttcagatggtactccag	6069
Db	2187	CTGGCCTTTGCCACCGAGCAGGACCCACCATCAGATGCTAAGATGCTACAGATGCTTCAG	2246
QY	6070	ggatctgttagycaccacagtgaaatcaggggaccttggaaagttgcccagggtttctgtct	6129
Db	2247	GGCTCTGTAGGGCCACCGTGAACCAAGGTCCTCTGGAGGTGGCCAGGTGTTTTCAGCA	2306
QY	6130	gaaataacctagtgaccacaaagctcttcagacatcataataaactgcgactcgtttaaa	6189
Db	2307	GAGATCCCGGAAGACCCCAAGCTTCTCCGGCATGCACAATAATTGCGGCTCTGCTTCAAG	2366
QY	6190	gatttactaaaggtgtgaagatgaccttaagaaaaataagagcttaattgggccggtt	6249
Db	2367	GACTTCTGCAAGAANAATGTGAGGATGCGCTGCGGAAANAATAAGCCCTGATGGGCCGAC	2426
QY	6250	caaaaggagtatacaaggygaattg--ggaaactatacttcgccttaagagggccctacag	6307
Db	2427	CAGAAGGAGTACCACCGGTAGCTGGAGCGCACTACTGCCGCTCGGGAGGCTCTGCAG	2486
QY	6308	ccctagatcacagaaagttccctcagttatccaagccagtatgtctgtccctgcacag	6367
Db	2487	CCCTGTCTTACCCAGCGCGCTGCCCAAGCTGATGGCACCACCC--ACCCGGCCTCAG	2542
QY	6368	agattccctcagtcgaatgaagccttcgcaaaatgagatctctaaac	6412
Db	2543	GAACTCTCTTGAACAGAGCAAGTTTCCGAAAGGACAGACCTCTGAGC	2587
RESULT 15			
AX174571			
LOCUS			
DEFINITION Sequence 3 from Patent WO0142296.			
ACCESSION AX174571			
VERSION AX174571.1 GI:14625747			
KEYWORDS			
SOURCE			
ORGANISM human.			
REFERENCE			
AUTHORS Lu,P., Garman,J.D. and Candia,A.F.			
TITLE Clasp-5 transmembrane protein			
JOURNAL Patent: WO 0142296-A 3 14-JUN-2001;			
Arbor Vita Corporation (US)			
FEATURES			
source			
CDS			
/codon_start=1			
/protein_id="CAC43926.1"			
/db_xref="GI:14625748"			
/translation="STMNADTAPTSPCPSISSQNSSSCSFQDOKIASMFDRTSRVPA			
SSTSSPGLFTFLAALDAEGEGISEVQRKAVSAIHSLSHDLDPRCVKPEVRVKA			
ALYLPVGLIIDLAPQLCDFTVADTRRYRTSGSDEDEGAGAITONVALAIAGNNFL			
KTSGLVLSLPYKQYMLNADTRNLICFLIMKNADQSLRKWIADLPSTQNLRL			
DLFICVLCFEYKQSSDQSVSTQYQKSDVKARLEAEALRGEGARGEEMRRRAGN			
DRFPGLENLRWKKEQTHWRQANEXLDTKAELDQELISGNLATEAHLIILDMQENI			
IQASSALDKDSLGGVLRVLVNSLNCDSSTYLTHTCFATRLIAKFGDLFEEVEE			
QCFDLCHQYVLAHCCSSMDVTRSQACATLYLLMRFSGATSNFARVKQMVTMSLASLVG			
RAPDFNEEHLRSLRTIILAYSEEDTAMQMTPEPTQVEELICNLNSILYDTVKMREFOE			
DPEMLDLMYRIAKSYQASPDRLRLTWLQNMAEKHKKCYTEAAMCLVHAALVAEYL			
SMLEDHSTLPVGSVSFYQNTSSNVLEESVSEDTLSPDEDEGAGQYFTESGLVLEQ			
AAELFSTGDIYETVNEVYKLVIPILAEHREFRKLTLTHSKLQRAFDISIVNKHKRMFG			

TYFRVGEFFGSKEGDDLDEQEEFVYKKEPATLKLPEISHRLEAFYGGCGAEFVEVIKIDST
VDKTKLDPNKAAYIQITFEVPEYFDEYEMKDRVTYFEKENFNLRREMYTTPFTLEGRPGE
LHEQYRRNTVLGTMHAFPIKTRISVIOKEEFVLTPIEVAIEDMKKTLQALVAINOE
PPDAKMLQAVLQSGSVGATVNOGPLEVAOVFLAEIPADPKLYRHNNKRLCEKEFIMRC
GEAVEKNKRLLITADQREYQOELKKNINLKNLRPMIERKIPELYPPIFRVESQKRDS
FHRSSFRCETQLSQGS"

BASE COUNT	1161	a	898	c	957	g	1010	t
ORIGIN								

OM of: US-09-737-246-2 to: GenEmbl:* out_format : pfs
Date: Jul 16, 2002 2:46 PM
About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09737246/runat_16072002_075430_26464/app_query.fasta_1.2180
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -GAPOP=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
-GAPOP=4.500 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
-FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blousum62
-TRANS=human40.cdi -LIST=45 -DOCCALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09737246_@CGN1_1_6982 -NCPU=6 -ICPU=3 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLPHY -WAIT -THREADS=1

Search information block:
Query: US-09-737-246-2
Query length: 2090
Database: GenEmbl:*
Database sequences: 1797656
Database length: 1873333701
Search time (sec): 5569.590000

score_list:					Strd Orig		ZScore	EScore	len	Documentation
Sequence					+ 10815.00	12925.95	0.0		6828	AX172874 Sequence 1 from Pat
gb_pat:AX172874					+ 7136.50	8519.28	0.0		6372	AX173175 Sequence 1 from Paten
gb_pr:AX173175					+ 6436.50	7685.07	0.0		4311	AB051558 Homo sapiens mRNA for
gb_pat:AB051558					+ 6171.50	7367.86	0.0		4200	AX172980 Sequence 107 from Pat
gb_pat:AX172980					+ 6157.50	7351.24	0.0		4144	AX172880 Sequence 7 from Paten
gb_pat:AX172880					+ 6157.50	7351.24	0.0		4144	AX172932 Sequence 59 from Paten
gb_pat:AX172932					+ 5785.50	6899.11	0.0		7215	AX174569 Sequence 1 from Paten
gb_pat:AX174569					+ 5472.50	6528.59	0.0		4886	AB037816 Homo sapiens mRNA for
gb_pr:AB037816					- 4986.50	5905.47	0.0		167215	AY016023 Spherooides nepheli
gb_ov:AY016023					- 4700.00	5571.69	5.5e-302		73515	AC017375 Drosophila melanoga
gb_htg:AC017375					+ 4690.00	5550.25	8.6e-301		166626	AC008318 Drosophila melano
gb_in:AC008318					+ 4690.00	5543.40	2.1e-300		301639	AE003590 Drosophila melano
gb_in:AE003590					+ 4674.50	5576.89	2.8e-302		3327	AK055905 Homo sapiens cDNA FLJ
gb_pr:AK055905					+ 4626.50	5519.63	4.3e-299		3253	AX172978 Sequence 105 from Pat
gb_pat:AX172978					+ 3528.50	4200.10	1.4e-225		4577	AK024436 Homo sapiens mRNA for
gb_pr:AK024436					+ 3235.50	3856.24	1.9e-206		2452	AK055493 Homo sapiens cDNA FLJ
gb_pr:AK055493					+ 3207.00	3820.46	1.9e-204		2825	BC008335 Homo sapiens, clone I
gb_pr:BC008335					+ 3122.50	3720.14	7.4e-199		2607	AK055401 Homo sapiens cDNA FLJ
gb_pr:AK055401					+ 3025.00	3598.30	4.5e-192		4026	AX174571 Sequence 3 from Paten
gb_pat:AX174571					+ 3025.00	3598.30	4.5e-192		4027	AX174664 Sequence 96 from Paté
gb_pat:AX174664					+ 2552.50	3039.41	6.1e-161		2149	AX173233 Sequence 59 from Paté
gb_pat:AX173233					+ 2480.50	2940.44	2.0e-155		6454	AX173022 Sequence 1 from Paten
gb_pat:AX173022					+ 2345.50	2776.95	2.6e-146		7506	AX255048 Sequence 7 from Paten
gb_pat:AX255048					+ 2273.50	2706.55	2.1e-142		1900	BC021414 Mus musculus, clone I
gb_ro:BC021414					+ 2245.50	2672.32	1.7e-140		2014	AX174662 Sequence 94 from Paté
gb_pat:AX174662					+ 2071.50	2459.62	1.2e-128		2904	BC019102 Homo sapiens, Similar
gb_pr:BC019102					- 2063.00	2400.26	2.4e-125		205035	AL049870 Human chromosome 14
gb_pr:CNS00005					+ 1929.50	2281.36	1.0e-118		5862	AB028981 Homo sapiens mRNA for
gb_pr:AB028981					- 1900.00	2224.16	1.6e-115		38886	U41543 Caenorhabditis elegans
gb_in:CEL46H5					+ 1780.00	2065.59	1.1e-106		139887	AT160233 Human chromosome 14
gb_pr:CNS01RGX					+ 1632.50	1928.84	4.4e-99		4391	AX173028 Sequence 7 from Paten
gb_pat:AX173028					+ 1615.50	1908.47	6.0e-98		4393	AX173118 Sequence 97 from Paté
gb_pat:AX173118					- 1609.00	1859.20	3.8e-95		159350	AC008354 Drosophila melanoga
gb_in:AC008354					- 1609.00	1853.67	6.8e-95		257224	AE003617 Drosophila melano
gb_in:AE003617					+ 1511.00	1784.64	4.8e-91		3899	AX058220 Sequence 90 from Paté
gb_pat:AX058220					+ 1484.50	1754.23	2.4e-89		3472	AX173120 Sequence 99 from Paté
gb_pat:AX173120					+ 1328.50	1568.43	5.3e-79		3152	AK056684 Homo sapiens cDNA FLJ
gb_pr:AK056684					+ 1183.00	1396.27	2.0e-69		2610	AX058192 Sequence 62 from Paté
gb_pat:AX058192					+ 1120.50	1311.97	1.0e-64		5896	AF465831 Arabidopsis thaliana
gb_pl:AF465831					+ 1111.00	1318.58	4.3e-65		1242	AF194407 Homo sapiens unknown
gb_pr:AF194407										

gb_pr:AK054649	+ 1097.00	1294.70	9.3e-64	2299	AK054649 Homo sapiens cDNA
gb_ro:RNTRG	+ 993.00	1166.17	1.3e-56	3227	X68101 R. norvegicus trg mRNA
gb_htg:AC020073	- 953.00	1082.75	5.9e-52	69694	AC020073 Drosophila melano
gb_ro:BC009134	+ 920.00	1079.55	9.0e-52	2998	BC009134 Mus musculus, Simi
gb_pr:BC015018	+ 901.00	1059.30	1.2e-50	2413	BC015018 Homo sapiens, clon

seq_name: gb_pat:AX172874

seq_documentation_block:

LOCUS AX172874 6828 bp DNA linear PAT 03-JUL-2001
DEFINITION Sequence 1 from Patent WO0142297.
ACCESSION AX172874
VERSION AX172874.1 GI:14597911
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 6828)
AUTHORS Lu, P., Garman, J.D. and Candia, A.F.
TITLE Clasp-3 transmembrane protein
JOURNAL Patent: WO 0142297-A 1 14-JUN-2001;
Arbor Vita Corporation (US)
FEATURES
source Location/Qualifiers
1..6828
/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="full-length human CLASP-3 cDNA"
22..6294
/note="human CLASP-3"
/codon_start=1
/protein_id="CAC43726.1"
/db_xref="GI:14597912"
/translation="MAERAAFAQKISRVAAEVRKQISGYSGSPQLKLNIVGNIS
HTTVPLEAVDPVDELYLITHPLAVDGGLRDLEFPDDIEVYSPRCRTLVSA
VPESEMDPHVRDCIRSYTEDWAIYIRKHLIGTFENPTLDQKREKQGLPKQVFS
DEAPDGNISYQDDQDDLRKMSIDTPRGWACSFIDLKNSLPDALLPNLDRTPNEE
IDRQNDQKRSNRHKEFLALHSPDEEPIERLSVPDIPKEHFGQLVKCLSKFEI
EIEPIFASIALYDVEKKEIKISEENGLINLEKLNLEGNHDGHRNSLASIYHFRLP
YPSQDVFLAIKLEKYLQGDIGCEAEFPMFEKADATKNKELEKKSQADFQCRLG
KYRPFAMTAIHMNIYSSAGSLERDSTEVEISTGERKGSWSERRNSIVGRSLERT
TSGDACNLTSFRPATLVTNFEKQEGDRLSDEDLYKFLADMRBSSVLRRLPITAQ
LKIDISPAENPHYCLTPELLQVKLYPDSRYRPTREILEFPARDVYVNTTYRNLYI
YPQSLFANRQGSARNITVVKQFMGYEDPSNAMPVIFGKSSCSSESKAAYTAVYHNR
SPDFHEIKVKLPALITDHHLLFTFYHVSQCKQNTPLETPVGVTWIPMLQNGRLKT
GQCLPVSLEKPPQAVSVLSPPEVPLPGMKWVNDHKGVFNEVVAVSSITHQDPYLDKF
FALVNALDEHLPVARGDMRIMENNLLENELKSSISALNSSQLPEVYRFLHLIDKLIL
LVIRPPIAGQIVNLGQASFEAMASIIINRHKNLGNHDGHRNSLASIYHFRLP
NTPYNSSSPGPGGLGSGSVHYATMARSAVPRASLNLNRSRLSNSNDPISGTPSPDE
VRSIIIGSKIDRSNWMVTGPKAPWGSNPSAESTQAMDRSCNRMSHTETSSFL
QTLTGLPTKCLFHEELALQWVVCSGSVRESALQAWFFELMVKSMVHLLYFNCKLE
APRKSREFPERFMDIALAVSTIASDIVSRFQDTEMERLNTSLAFLINDLLSVMDRG
FVFSLIKSCYKQVSSKLYSLPNPSVLVSLRDLFLRICSHHYVTLNPLCSLITPPAS
PSPVSSATSQSGSFSTNVQDQKIANMFELSVPRQOHLAGVLTELAVIDLPPAEG
LFGLLKKVINMVLNLSHSDSDPRYSDPQIKARVAMLVPLIGIIMEVTPQLYDFET
HNQGRPICIAITDYESGSGMSIQTVAMAIAGTSVPQLTRPGSFLLTSTSGRQHTTF
SAESSRSLICLWVLEKNADETVLQKFTDLSVLQNLRLDLLYLCVSCFEYKGVF
ERMNSLTFKKSMDRAKLEAEILGSIGARQEMVRRSGQLERSPGSAGSOENLRWR
KDMTHWRQNTKELDKRAEIEHEALIDGNLAEANLILDTLEIVQTVSYTESKESI
LGGLVKVLHSMACQSAVYLQHCFAQRAIVLSKPELLEFEETEQCADLCRLLRHC
SSSIGTIRSHPSASLYLLMRQNFIEGNNFARKMQVPMSSLSIVGTSGNFNEFLRS
LKITLYAEEDLELRETFPPDQVODLVFNHLHILSDTVKMKEHQEDPMLIDLMYRIA
KGYQTSPELRLQNNAGKHSERSNHAEOCLVSAALVAEYLSMLEDKRYLPVGCY
TEQNISSNVLLESAVSDVVSDESGKXTESGLVGLLEQAAFSMAGMYEAV
NEVYKVLPIPIHEANRAKKLSTIHGKLOEAPSKIVHQSTGWMRMFGYFRVGTGKF
IQDLEQEFVKEPALITKLAIESIRLGEFYGERFGEDVLEIKDSNPDKLQKPKITAY
IQTIVVEPYFDTYEMKDRITTFYFDKNILTRPIEVALIEDMQKTOELAFATHQDADPKMLQWVLQ
GSGVTYNQGPLVAVQVFLSEIPSPDKLFRHNNKLRLCFKDFTKCEDALRKNNKSLIG
PVQKEYORELGLSSP"

CDS

BASE COUNT
ORIGIN

2082 a 1384 c 1432 g 1930 t

```

alignment_scores:
    Quality: 10815.00
    Ratio: 5.175
    Percent Similarity: 100.000
    Percent Identity: 100.000
alignment_block:
    US-09-737-246-2 x AX172874 ..
Align seg 1/1 to: AX172874 from: 1 to: 6828

1  MetaIaGlUaRgaRgaLaPheAlaGlInLysIleSerArGthrValAlaI 17
   |||||||
22  ATGCCGAGCGCCGCCCTTCGCCAGAGATCAGCAAGCGGTGGCAGC 71
   |||||||
17  aGlUValaRgLySgInIleSerGlyGInTyRserGlySerProGlnleuL 34
   |||||||
72  CGAAGTTAGGAGCAGATCTCCGACAATATAGTGTTCTCCCAACTGC 121
   |||||||
34  eulYsAsnleuAsnIleValGlyAsnIleSerHisThrThrValPro 50
   |||||||
122 TCAAAAACCTTAATATGTTGGCAATATATCCCATCACACCAGTGGCC 171
   |||||||
51  leuThrGlUalValAspProValAspleuGluAspTyRleuIleThrHi 67
   |||||||
172 CTTACCGAAGCAGATGATCCAGTGGATTGGAGAAGATTACCTCATTTACTCA 221
   |||||||
67  sProleuAlaValAspSerGlyProleuArGAspleuIleGluPhePro 84
   |||||||
222 TCCTTTGGCTGGATTCCTGGCCCTTTACGGGATTGTGATTTGAATTTCCCTC 271
   |||||||
84  roAspAspIleGluValValTyRserProArGAspCysArGThrleuVal 100
   |||||||
272 CAGATGATATGTGAAGTTGTTATAGTCCCTGGGACTGCAGAACTCTGT 321
   |||||||
101 SerAlaValProGluGluSerGluMetAspProHisValArGAspCysII 117
   |||||||
322 TCAGCTGTACCTGAAGAAAGTGAATGATGCCATGTTAGAGACTGTAT 371
   |||||||
117 eArGSerTyRThrGluAspTrpAlaIleValIleArGlySTyRHisLysL 134
   |||||||
372 AAGAAGTTATACAGAAGACTGGGCAATTGTCATCAGAAAATATCATTAAT 421
   |||||||
134 euGlyThrGlyPheAsnProAsnThrleuAspLysGlnLysGluArGln 150
   |||||||
422 TGGGAACAGGATTTAATCCCAATACATGATGATAAACAGAAAGAAAGCAA 471
   |||||||
151 lysGlyLeuProLysGlnValPheGluSerAspGluAlaProAspGlyAs 167
   |||||||
472 AAAGGTTTGCCAAACAAGTTTTTGAAATCGATGAAGCTCCAGATGGCAA 521
   |||||||
167 nSerTyRglnAspAspGlnAspAspleuLysArGArGserMetSerIleA 184
   |||||||
522 CAGCTACCAGAGATGATCAAGATGACCTTAAAGAAGCTTCAATGTCAATAG 571
   |||||||
184 sPasPThrProArGlySerTrpAlaCysSerIlePheAspleuLysAsn 200
   |||||||
572 ATGATACCCCAAGGGGTAGCTGGGCCGTGATATCTTGACTTGAAAAAT 621
   |||||||
201 SerleuProAspAlaIleuLeuProAsnleuLeuAspArGThrProAsnGI 217
   |||||||
622 TCACCTTCCTGATGCTTTGCTTCCCAATTAATTGATCGAAGCTCCAATGA 671
   |||||||
217 uGlUleAspArGlnAsnAspAspGlnArGlySserAsnArGHisLysG 234
   |||||||
672 AGAATAGACCGTCAGATGATGACCAAGAAATCAAAACCGTCACAAAG 721
   |||||||
234 luLeuPheAlaIleuHisProSerProAspGluGluProIleGluArG 250
   |||||||
722 AACTTTTGCTTGATCCATCACCAGATGAGGAAGAACCAATAGAACGG 771
   |||||||
251 leuSerValProAspIleProLysGlnHisPheGlyGlnArGleuLeuVa 267
   |||||||

```

772	CTTAGTGTTCCTGATATATACCCAAAGAACATTTTGGTCAAGACCTTCTTGT	821
267	llysCysteSerLeuLysPheGluIleGluIleGluProIlePheAla	284
822	AAAATGCTTATCACTCAAGTTGAAATTGAAATTGAACCCATTTTGGCAA	871
284	erLeuAlaIleTyrAspValLysGluLysLysLysIleSerGluAsnPhe	300
872	GTTTGGCTTATATGATGTCAAGCAAAAGAAAAGATTTCAGAAACTTT	921
301	TyrPheAspLeuAsnSerGluGlnMetLysGlyLeuLeuArgProHisVa	317
922	TATTTTGACCTTAATTCTGAGCAGATGAAGGGTTGTTACGTCACATGT	971
317	lProProAlaAlaIleThrThrLeuAlaArgSerAlaIlePheSerIleT	334
972	ACCACCTGCTGCCATTACTACCTGGCAAGATCAGCAATTTTCTATCA	1021
334	hTyrProSerGlnAspValPheLeuValIleLysLeuGluLysValLeu	350
1022	CTTATCTTCCCAAGATGTTTCTTGTATTAAGCTAGAAAAAGTCTTA	1071
351	GlnGlnGlyAspIleGlyGluCysAlaGluProTyrMetIlePheLysG	367
1072	CAGCAAGAGACATTGGAGAGTGTGCAGAACCATATATGATTTTCAAGA	1121
367	uAlaAspAlaThrLysAsnLysGluLysLeuGluLysLeuLysSerGlnA	384
1122	AGCAGATGCCACCAAGATAAAGAAAACTGGAGAACTGAAGAGTCAAG	1171
384	lAspGlnPheCysGlnArgLeuGlyLysTyrArgMetProPheAlaTyr	400
1172	CAGATCAGTTTGGCCAAAGACTTGGGAAATATCCGATGCCCTTTGCTTG	1221
401	ThrAlaIleHisLeuMetAsnIleValSerSerAlaGlySerLeuGluAr	417
1222	ACTGCAATCCATTTAATGAATATTGTTAGCAGTGTCTGGAGTTTGGAAAG	1271
417	gAspSerThrGluValGluIleSerThrGlyGluArgLysGlySerTrp	434
1272	AGATTCTACAGAAGTAGAAATCAGTACTGGAGAACGAAAAAGGCTTGTGT	1321
434	erGluArgArgAsnSerSerIleValGlyArgArgSerLeuGluArgThr	450
1322	CAGAGAGAGGAATTTAGTATGTTGGCAGACGATCACTTGAAGGACA	1371
451	ThrSerGlyAspAspAlaCysAsnLeuThrSerPheArgProAlaThrLe	467
1372	ACAAGTGGAGATGATGCTGTGACTTGACGAGCTTTCGACCAGCTACTCT	1421
467	uThrValThrAsnPhePheLysGlnGluGlyAspArgLeuSerAspGluA	484
1422	CACAGTGACAAATTTTTTAAAGCAGGAAGAGACCCGCTTAAGTATGAAG	1471
484	sPleuTyrLysPheLeuAlaAspMetArgArgProSerSerValLeuArg	500
1472	ATCTCTACAAATTCCTTGTGATATGAGAAAGGCCATCTTCTGTCTTACGG	1521
501	ArgLeuArgProIleThrAlaGlnLeuLysIleAspIleSerProAlaArg	517
1522	CGACTAAGACCTATTACAGCTCAGCTCAAGATAGACATTTCTCCCGCAC	1571
517	oGluAsnProHisTyrCysLeuThrProGluLeuLeuGlnValLysLeuT	534
1572	TGAAAATCCCCCATTTATTCGCCTAACCTCCGAGCTGCTTCAAGTGAAGCTTT	1621
534	yrProAspSerArgValArgProThrArgGluIleLeuGluPheProAla	550
1622	ACCCTGACAGTAGAGTTAGACCTACAGAGAAATCTTAGAGTTTCCCGCA	1671
551	ArgAspValTyrValProAsnThrThrTyrArgAsnLeuLeuTyrIleTy	567
1672	AGGGATGTTTATGTTCCAAACACTACTTACAGAAATCTTCTACATATA	1721

567 rProGlnSerLeuAsnPhelaAsnArgGlnGlySerAlaArgAsnIleT 584
|||||
1722 CCCTCAGAGCTCTTAATTTTGCCAATCGTCAGAGTTCTGCTAGAAATATAA 1771
584 hrValLysValGlnPheMetTyrGlyLuAspProSerAsnAlaMetPro 600
|||||
1772 CAGTGAAAGTCCAGTTTATGTATGAGAGATCCAGCAATGCCATGCCG 1821
601 ValIlePheGlyLysSerSerCysSerGluPheSerLysGluAlaTyrTh 617
|||||
1822 GTAATCTTTGGTAAATCTAGCTGTTCAGAAATTTCAAGAGACCTATAC 1871
617 rAlaValValTyrHisAsnArgSerProAspPheHisGluGluIleLysV 634
|||||
1872 AGCCGTAGTATATCATTAACAGGCTCTCCTGATTTTCATGAAGAAATCAAG 1921
634 aLLysLeuProAlaThrLeuThrAspHisHisLeuLeuPheThrPhe 650
|||||
1922 TTAAGCTTCCCTGCTACTTTAACTGACCATCATCATCTGCTTTTACTTTT 1971
651 TyrHisValSerCysGlnGlnLysGlnAsnThrProLeuGluThrProVa 667
|||||
1972 TATCATGTAGTGTGTCAACAACAAATACTCCTCTTGAAACACACAGT 2021
667 lGlyTyrThrTriPleProMetLeuGlnAsnGlyArgLeuLysThrGlyG 684
|||||
2022 TGGATATACATGGATACCAATGCTTCAGAAATGGACGGTTGAAGACTGGCC 2071
684 lInPheCysLeuProValSerLeuGluLysProProGlnAlaTyrSerVal 700
|||||
2072 AGTTTGCTTGCCAGCTCTCATTTGGAAAAACCAACCACACAGCCTTATTCTGTA 2121
701 LeuSerProGluValProLeuProGlyMetLysTrpValAspAsnHisLy 717
|||||
2122 CTGTCTCCTGAGGTTCCCTTACTGTCGATGAAATGGGTAGATATACACAA 2171
717 sGlyValPheAsnValGluValAlaValSerSerIleHisThrGlnA 734
|||||
2172 AGGTGTTTTAATGTGAAGTGTGTGCTGTTCTGTATCCATACACAAAG 2221
734 sPrProTyrLeuAspLysPhePheAlaLeuValAsnAlaLeuAspGluHis 750
|||||
2222 ATCCTTATCTTGACAAATTTTGTCTGTGTCATGCTCTGGATGAACAC 2271
751 leuPheProValArgIleGlyAspMetArgIleMetGluAsnAsnLeuG 767
|||||
2272 CTGTTCCCAGTCCGAATGGGGACATGCGAATCATGGAAAAATAACTTAGA 2321
767 uAsnGluLeuLysSerSerIleSerAlaLeuAsnSerSerGlnLeuGluP 784
|||||
2322 AAATGAATTGAAAGACAGATTTTCAGCACTGAATTCATCCCAAGCTGAAC 2371
784 roValValArgPheLeuHisLeuLeuLeuAspLysLeuIleLeuLeuVal 800
|||||
2372 CAGTGTCCGATTTCTTCATCTTCTGCTAGATAAAGTGAATGCTTTAGTT 2421
801 lLeArgProProValIleAlaGlyGlnIleValAsnLeuGlyGlnAlaSe 817
|||||
2422 ATTAGACCTCCTGTCATGTCGCCCAAATAGTTAAACCTAGGTCGAAGCATC 2471
817 rPheGluAlaMetAlaSerIleIleAsnArgLeuHisLysAsnLeuGluG 834
|||||
2472 TTTTGAAGCCATGGCATCAATATAAATCGACTTCACAAAAAACTTGAAG 2521
834 lYAsnHisAspGlnHisGlyArgAsnSerLeuLeuAlaSerTyrIleHis 850
|||||
2522 GAAATCATGACACGATGGCAGAAACAGCCTTCTTGCAATCATATATTATCAT 2571
851 TyrValPheArgLeuProAsnThrTyrProAsnSerSerSerProGlyPr 867
|||||
2572 TATGTTTTCCGCCCTACCAAAATACTTACCCTTAATTCATCATCACCAAGTCC 2621

867 oGlyGlyLeuGlyGlySerValHisTyrAlaThrMetAlaArgSerAlaV 884
|||||
2622 TGGGGGTTTGGGAGGATCAGTGCATTTATGCCACAATGGCTAGATCTGCCG 2671
884 aLArgProAlaSerLeuAsnLeuAsnArgSerArgSerLeuSerAsnSer 900
|||||
2672 TGAGACCTGCAAGCCTTAATTTAAATCGTTCGAGAGCCTTAGTAATAGC 2721
901 AsnProAspIleSerGlyThrProThrSerProAspAspGluValArgSe 917
|||||
2722 AATCCAGATATATCTGGGACTCCACGTCCACAGATGATGAAGTTGATC 2771
917 rIleIleGlySerLysGlyLeuAspArgSerAsnSerTrpValAsnThrg 934
|||||
2772 AATCATCGGGAGTAAAGGTTTAGATCGCTCCAAATTCCTGGGTTAACACTG 2821
934 lYgLYProLysAlaAlaProTriPglySerAsnProSerProSerAlaGlu 950
|||||
2822 GTGGTCCAAAGCTGCCCATGGGATGCCAACCCAGTCAGTGACAGAA 2871
951 SerThrGlnAlaMetAspArgSerCysAsnArgMetSerSerHisThrGl 967
|||||
2872 TCAACACAGGCTATGATCGAAGTTGTATATCGTATGTCTTCGCACACAGA 2921
967 uThrSerSerPheLeuGlnThrLeuThrGlyArgLeuProThrLysLysL 984
|||||
2922 GACGTCAAGTTTCTTACAAACATTAACGGGACGCTTACCACACTAAAAAGC 2971
984 euPheHisGluGluLeuAlaLeuGlnTrpValValCysSerGlySerVal 1000
|||||
2972 TTTTTCACGAGGAGCTGGCTTTCAGTGGGTGTTGTCAGTGGCAGCGTT 3021
1001 ArgGluSerAlaLeuGlnGlnAlaTrpPhePhePheGluLeuMetValLy 1017
|||||
3022 CGGAATCAGCTTTGCAACAACAGCCTGCTCTTTTGAATTAATGTGTA 3071
1017 sSerMetValHisHisLeuTyrPheAsnAspLysLeuGluAlaProArgL 1034
|||||
3072 GAGCATGGTGCACCATTTATACTTTAATGATAAAGTTGAGGCTCCCAAGGA 3121
1034 ySerArgPheProGluArgPheMetAspIleAlaAlaLeuValSer 1050
|||||
3122 AAAGTCGTTTCCAGAACGTTTCATGATGACATTTGCAGCTCTTGTCAGC 3171
1051 ThrIleAlaSerAspIleValSerArgPheGlnLysAspThrGluMetVa 1067
|||||
3172 ACGATTGCTAGTATAGTTTTCACGATTTTCAGAAAGACACAGAAATGGT 3221
1067 lGluArgLeuAsnThrSerLeuAlaPhePheLeuAsnAspLeuLeuSerV 1084
|||||
3222 TGAGAGACTCATATACAAGCCTTGCAATCTTTCTCATGATGATGTTGTCG 3271
1084 alMetAspArgGlyPheValPheSerLeuIleLysSerCysTyrLysGln 1100
|||||
3272 TTATGACACAGAGATTTGTTTTTAGCCTTATAAAGTCTGCTATAAAGAG 3321
1101 ValSerSerLysLeuTyrSerLeuProAsnProSerValLeuValSerLe 1117
|||||
3322 GTGCTTCAAGCCTTTACTCATTTACCGAATCCAGTGTCTGTGTCCTT 3371
1117 uArgLeuAspPheLeuArgIleIleCysSerHisGluHisTyrValThrL 1134
|||||
3372 GAGGCTGGATTTTCTACGAATCATGTGATGAGACACTATGTTACAT 3421
1134 euAsnLeuProCysSerLeuLeuThrProAlaSerProSerProSer 1150
|||||
3422 TAAACTTACCCGACAGCTTACTTACCCACCTGCATCTCCATCACCTTCT 3471
1151 ValSerSerAlaThrSerGlnSerSerGlyPheSerThrAsnValGlnAs 1167
|||||
3472 GTTCTCTGCAACATCTCAGAGTTCGATTTTCTACGAATGTACAAGA 3521
1167 pGlnLysIleAlaAsnMetPheGluLeuSerValPropheArgGlnGlnH 1184

|||||
3522 CCAAAGATTGCAAAATATATGTTGAATTAATCCGTGCCTTTCGCCGCAACAGC 3571
1184 lStYrLeuAlaGlyLeuValLeuThrGluLeuAlaValIleLeuAspPro 1200
3572 ATTAATTGGCAGACTGTGTGTTAACAGAGCTGGCTGTCAATTATTAGACCTT 3621
1201 AspAlaGluGlyLeuPheGlyLeuHisLysLysValIleAsnMetValHis 1217
3622 GATGCTGAAGGACTGTTGGATTGCATAGAAGATCATCAATATGTTACA 3671
1217 sAsnLeuLeuSerSerHisAspSerAspProArgTyrSerAspProGlnI 1234
3672 CAATTTACTCTCCAGTCAACGACTCAGACCCGCGGTACTCTGACCTCACA 3721
1234 lelyAlaArgValAlaMetLeuTyrLeuProLeuIleGlyIleIleMet 1250
3722 TAAAGGCTCGAGTGGCCATGTTGATCTACCTCTGATTGGTATTCATG 3771
1251 GluThrValProGlnLeuTyrAspPheThrGluThrHisAsnGlnArgG 1267
3772 GAAACTGTACCTCAGCTGTATGATTTTACAGAAACTCACAATCAACGAGG 3821
1267 YArgProIleCysIleAlaThrAspAspTyrGluSerGluSerGlySerM 1284
3822 AAGACCAATTTGTAATAGCCACTGATGATTATGAAGTGAGAGCGGAAGTA 3871
1284 etIleSerGlnThrValAlaMetAlaIleAlaGlyThrSerValProGln 1300
3872 TGATTAAGCCAGACCGTTGCCATGGCAATCGCAGGACATCGGTCCCTCAA 3921
1301 LeuThrArgProGlySerPheLeuLeuThrSerThrSerGlyArgGlnHis 1317
3922 CTAAACAAGCGCTGGCAGATTTCCTCCTCAGCTCAACGAGTGGCAGGCAACA 3971
1317 sThrThrPheSerAlaGluSerSerArgSerLeuLeuIleCysLeuLeuT 1334
3972 CACTACCTTTTCAGCAGAATCAAGTCGAAGCCTTTTGATCTGTCTACTTT 4021
1334 rpValLeuLysAsnAlaAspGluThrValLeuGlnLysTrpPheThrAsp 1350
4022 GGGTTCCTCAAAAATGCAGATGAACACAGTTCTACAGAAGTGTTTACAGAT 4071
1351 LeuSerValLeuGlnLeuAsnArgLeuLeuAspLeuLeuTyrLeuCysVa 1367
4072 CTCTCAGTCTTGACAGCTAAACCGGCTATTAGATCTGCTTATCTCTGTGT 4121
1367 lSerCysPheGluTyrLysGlyLysLysValPheGluArgMetAsnSerL 1384
4122 GTCTTGCTTTGAGTATTAAGGAAAAAGTGTGTAACGAATGAATAGCT 4171
1384 euThrPheLysLysSerLysAspMetArgAlaLysLeuGluGluAlaIle 1400
4172 TGACCTTTTAAGAAATCAAAAAGACATGAGCAAAAGCTTGAAGAAGCTATT 4221
1401 LeuGlySerIleGlyAlaArgGlnLumetValArgArgSerArgGlyG 1417
4222 CTTGGAGCATAGGTGCCAGGCAAGAAATGGTACGGCGAAGCCGAGGACA 4271
1417 nLeuGluArgSerProSerGlySerAlaPheGlySerGlnGluAsnLeuA 1434
4272 GCTCGAGAGAAAGCCCATCTGGAAGTGCCTTTGGAAGTCAAGAAAAATTTGA 4321
1434 rgtTrpArgLysAspMetThrHisTrpArgGlnAsnThrGluLysLeuAsp 1450
4322 GGTGAGAGAAAGATATGACTCAGTCGCGTCAAAAACACAGAGAAAGCTTGAC 4371
1451 LysSerArgAlaGluIleGluHisGluAlaLeuIleAspGlyAsnLeuAl 1467
4372 AAATCAAGAGACAGATTGAACACAGAACACTGATTGATGGAACCTGGC 4421
1467 aThrGluAlaAsnLeuIleIleLeuAspThrLeuGluIleValValGlnT 1484
|||||

4422 TACAGAAGCAAAACCTAAATCATTTTAGATACATTAGAGATTGTTGTCAGA 4471
1484 hrValSerValThrGluSerLysGluSerIleLeuGlyGlyValLeuLys 1500
4472 CCGTTTCTGTAAACGGAATCCAAAGAGAGCATTTCTTGGAGTGCTTAA 4521
1501 ValLeuLeuHisSerMetAlaCysAsnGlnSerAlaValTyrLeuGlnHis 1517
4522 GTGCTACTACACAGCATGGCCCTGTAACCAAGATGCAGTTTATCTACACA 4571
1517 sCysPheAlaThrGlnArgAlaLeuValSerLysPheProGluLeuLeuP 1534
4572 CTGTTTTCCTACACAGAGAGCCCTTGTTCAAGTTTCTCGAAGCTTTAT 4621
1534 heGluGluGluThrGluGlnCysAlaAspLeuCysLeuArgLeuLeuArg 1550
4622 TTGAAGAAGAGACAGACAGTGTGCTGATTATTATGCCCTCAGGCTTCGCA 4671
1551 HisCysSerSerSerIleGlyThrIleArgSerHisProSerAlaSerLe 1567
4672 CACTGTAGCAGTAGCATCGGTACAAATACGGTCAACACCCAGTGCTCCCT 4721
1567 uTyrLeuLeuMetArgGlnAsnPheGluIleGlyAsnAsnPheAlaArgV 1584
4722 TTACCTACTAATGAGGCAAAACTTTGAGATTGGGAATAACTTTGCCAGGG 4771
1584 alLysMetGlnValPrometSerLeuSerSerLeuValGlyThrSerGln 1600
4772 TTAATAATGCAGGTACCAATGTCACTATCTCTTGTTGGTGCCACATCTCAG 4821
1601 AsnPheAsnGluGluPheLeuArgArgSerLeuLysThrIleLeuThrTy 1617
4822 AATTTTAATGAAGAAATCTTAAGACGTTCTTAAGACTATATATTGACATA 4871
1617 rAlaGluGluAspLeuGluLeuArgGluThrThrPheProAspGlnValG 1634
4872 TGCTGAAGAAGATCTGGAATTGAGGGAACAACATTTCTCGATCAGGTCC 4921
1634 lAspLeuValPheAsnLeuHisMetIleLeuSerAspThrValLysMet 1650
4922 AGCATCTGGTTTCAATCTCCATATGATGATCTTCTGATACTGTGAATAATG 4971
1651 LysGluHisGlnGluAspProGluMetLeuIleAspLeuMetTyrArgIl 1667
4972 AAGGAACACCAGAGAGATCCTGAATGTGTGATGATCTAATGTACAGAAAT 5021
1667 eaLalysGlyTyrGlnThrSerProGluArgLeuThrTrpLeuGlnAsnM 1684
5022 TGCCAAAGGTTACCAGACCTCTCCAGAGCGAATTGACCTGGTTGCAGACA 5071
1684 etAlaGlyLysHisSerGluArgSerAsnHisAlaGluAlaAlaGlnCys 1700
5072 TGGCAGGCAAGCACTCAGAACGAAAGCAATCATGCTGAAGCTGCACAGTGT 5121
1701 LeuValHisSerAlaAlaLeuValAlaGluTyrLeuSerMetLeuGluAs 1717
5122 CTAGTCACACTCAGCAGCACTTGTGCTGAATATTTGAGCATGCTGGAGGA 5171
1717 pArgLysTyrLeuProValGlyCysValThrPheGlnAsnIleSerSerA 1734
5172 CCGGAATATCTTCCCTGTGGGATGTGTAAACATTTCAAGATATTTCATCTA 5221
1734 snValLeuGluGluSerAlaValSerAspAspValValSerProAspGlu 1750
5222 ATGTTTGAAGAAATCTGCGGTCTCAGATGATGTGTATCTCCAGATGAA 5271
1751 GluGlyIleCysSerGlyLysTyrPheThrGluSerGlyLeuValGlyLe 1767
5272 GAAGGTATCTGCTGTGAAAAATACTTTACTGAGTCAGGACTTGTGGGATT 5321
1767 uLeuGluGlnAlaAlaLaserPheSerMetAlaGlyMetTyrGluAlay 1784
5322 ACTGGAACAAGCAGCTGCTCTCTCTATGGCTGGCATGTATGAACGACG 5371
|||||

1784 alaIngIuValTyrLysValleuIleProIleHisGluAlaasnArgasp 1800
|||||
5372 TTAATGAAGTTTACAAAGTACTTATTCCTATTCATGAAGCTAATCGGAT 5421
1801 AlaLysLysLeuSerThrIleHisGlyLysLeuGlnGluAlaPheSerLy 1817
|||||
5422 GCAAGAAACTATCCACAAATTCATGGTAACCTTCAGAAGCATTCAGCAA 5471
1817 sIleValHisGlnSerThrGlyTrrGluArgMetPheGlyThrTyrPhea 1834
|||||
5472 AATTGTTCAACAGAGTACTGGCTGGAGCGGATGTTGGCACCTATTTC 5521
1834 rGValGlyPheTyrGlyThrLysPheGlyAspLeuAspGluGlnGluPhe 1850
|||||
5522 GTGTTGGTTTTTATGGACCAAGTTCGGGATTTGGATGAACAAGAAATTT 5571
1851 ValTyrLysGluProAlaIleThrLysLeuAlaGluIleSerHisArgLe 1867
|||||
5572 GTTACAAAGAGCCTGCAATTAACCAACTTGACAGAGATATCTCACAGATT 5621
1867 uGluGlyPheTyrGlyGluArgPheGlyGluAspValValGluValIleL 1884
|||||
5622 GGAGGGATTTTACGGAGAAGATTTGGAGAGGATGTGGTTGAAGTATCA 5671
1884 ysAspSerAsnProValAspLysCysLysLeuAspProAsnLysAlaTyr 1900
|||||
5672 AAGACTTAATCTCTGTAGACAGTGAATATAGATCCTTAACAAGGCATAT 5721
1901 ILeGlnIleThrTyrValGluProTyrPheAspThrTyrGluMetLysAs 1917
|||||
5722 ATTCAAGATTACCTATGTGGAGCCATACTTTGACACATATGAGATGAAGA 5771
1917 pArgIleThrTyrPheAspLysAsnTyrAsnLeuArgArgPheMetTyrC 1934
|||||
5772 CAGAAACACCTATTTCCAGCAAAATTAACAATCTTCGCATTCATGTACT 5821
1934 ysThrProPheThrLeuAspGlyArgAlaHisGlyGluLeuHisGluGln 1950
|||||
5822 GTACACCTTTACTTTAGATGGCCGTGCCCATGGGAACTTCATGAACAA 5871
1951 PheLysArgLysThrIleLeuThrThrSerHisAlaPheProTyrIleLy 1967
|||||
5872 TTCAAAAGGAAGACCATTCGTGACTACGTCATCGCCTTCTTATATTAA 5921
1967 sThrArgValAsnValThrHisLysGluGluIleIleLeuThrProIleG 1984
|||||
5922 AACAAAGGTCATGTCTACTCATTAAGAAGAGATCATCTTAACCACAATTG 5971
1984 luValAlaIleGluAspMetGlnLysLysThrGlnGluLeuAlaPheAla 2000
|||||
5972 AAGTGTCTATTGAGACATGCAAGAAAAAGACACAGAGAGTTGGCATTTGCA 6021
2001 ThrHisGlnAspProAlaAspProLysMetLeuGlnMetValleuGlnG 2017
|||||
6022 ACACATCAGGATCCCGCAGAGCCCAAAATGCTTCAGATGTACTCCAGGG 6071
2017 ySerValGlyThrThrValAsnGlnGlyProLeuGluValAlaGlnValP 2034
|||||
6072 ATCTGTAGGCACACACAGTGAATGAGGGCCTTTGGAAGTTGCCACGCTT 6121
2034 heLeuSerGluIleProSerAspProLysLeuPheArgHisAsnLys 2050
|||||
6122 TTCTGTCTGAATACTAGTAGACCAAAAGCTCTTCAGACATCATATAATAA 6171
2051 LeuArgLeuCysPheLysAspPheThrLysArgCysGluAspAlaLeuAr 2067
|||||
6172 CTGCGACTCTGCTTTAAAGATTTTACTAAAAAGGTGTGAAGATGCCCTTAAG 6221
2067 gLysAsnLysSerLeuIleGlyProValGlnLysGluTyrGlnArgGluL 2084
|||||
6222 AAAAAATAAGAGCTTAATTGGCGCGTTCAAAAAGAGATATCAAAAGGAAT 6271

2084 eugLyLysLeuSerSerPro 2090
|||||
6272 TGGGAAACTATCTTCGCCT 6291
seq_name: gb_pat:AX173175

seq_documentation_block: 6372 bp DNA linear PAT 03-JUL-2001
LOCUS AX173175
DEFINITION Sequence 1 from Patent WO0142295.
ACCESSION AX173175
VERSION AX173175.1 GI:14598036
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 6372)
AUTHORS Lu, P., Garman, J.D. and Candia, A.F.
TITLE Clasp-7 transmembrane protein
JOURNAL Patent: WO 0142295-A 1 14-JUN-2001;
Arbor Vita Corporation (US)

FEATURES
source location/Qualifiers
1..6372
/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="full length human CLASP-7 cDNA"
13..6156
/note="human CLASP-7"
/codon_start=1
/protein_id="CAC43731.1"
/db_xref="GI:14598037"
/translation="MAASERRAFAHKINRYAAAEVRKQVSRERSGSPHSRRCSSSLG
VPLTEVEVERPLDFEDVLLSRPPDAEGRPLRDLYEFPADDELILLQPREGRTTEPGIPKD
EKLDQAQRAAVEMYIEDWTVHRRYQYLSAAYSPVTTDQREKQKQPROQFEODASG
DERSGPEDSNDNRGSGSPEDTPRSSGASSIFDLRLNLAADSLPSLLERAAPEDVDR
NETLRQHRPPALTLTYPADEDEAVERCSRPPEHFGQRIYVCLSLKEIEIEIP
IEGILALYDVREKKITSENFYFDLNSDSMKGLIRAHGTHPAISTLARSAIFSVTYPS
DIFLVITLKEKVLQOQDISECEPYMVLKEVDTAKNKELEKELRLAEOECTRLGRYRM
PFAWTAVHLANIVSSAGQDLDRSDSEGERRPAMTDRRRRGPODRASGDDACSFSGFR
PATLTVTNEFKQEAERLSDLEKFLADMRRSSILRLRPVTAQIKIDISPAPENPH
FCLSPELHIKPYDPGRGRPTKEILEFPAREVYAPHTSYRNLLYYPHSLNESSRQGS
VRNLAVRQYMTGEDPSQALPIIFGKSSCSEFTREAFTPVYVHNKSPFEYEELHLHP
ACVTENHLLFTFYNVSCQPRPGTALETPVGTWIPLQHGRLRTGFCLPVSDVDP
PSYSVLTLPDVALPGMRVWDGHKGVFSVELTAVSSVHPQDPYLDKFTLVHVEGAFP
FRUKDTVLSEGNTVEQELRASLALRLASPEPLVAESHVLDKLVRLVTRPIISGOIV
NLGRGAFEMAHVSVLVHRSLEAAQDARGHPQLAAVYVHYAFRLPGTEPSLPDGAAPV
TYQAATLARGSGRPASLYLARSKSISSSNPDLAVAPGSVDDEVSRILASKLHEELAL
QWVSVSSAVREALIQHAMEFFQLMVKSMALHLLGLQRLDTPRKLRFGREFLDITALV
GSVGLFVETRVHKKDVELAEHLNLSIAFLSDLLSLVDRGFVFSLVRAHYKQVATRILOS
SPNPALTLTLMETRIILCSHEHYVTNLNLPCCPLSPSPSPSPSVSTQSSTFSSQA
PDPKVTSMFELSGPFRQOHFLAGLLTELALALEPEAGAFLLHKKAISAVHSLLCGH
DIDPRYAEATVKARVAEDLYPLLSIARDTLPLRHDFEAGPGQORSRLASMLDSDEEGE
DIAGTINPSVAMAIAGBPPLAGPSRASISQGPPTASRAGCALSAESSRLLACVWLVK
NTEPALIQRWATDITLTPQLGRLLDLILYLCLAAFEYKGAFAFERINSITFKKSLIDMKAR
LEEAALGTIGARQEMVARSRSRSPFGNPENVRWRKSVTHWKQTSDRVDKTKDEMEHEA
LVEGNLATEASLVVLDLTLEIIVQTVMLSEARESVLGAVLKVVLYSLGSQSAFLIOHG
LATQRALVSKFPELLFEEDTELCADECLRLLRHCGSRISTIRTHASASLYLLMRONE
IGHNEFARVKMOVTMSLSLVGTQNFSEHLSRSIKTILTYAEEDMGRLSDTFAEOVO
DIMENLHMILTDITYKMKHEQDEPEMLIDLMYRIARGYQGSPLRLTLQNMAGKHAEL
GNHAEAAQCMVHAALVAEYLALLLEDORHLPPGCVSEONISSNVLEESAISDDILSPD
BEGFSGKHFTELGLLEQAAGYFTMGGLYEAENEVYKNLILPILEAHRDYKKLAAY
HGKIQEAFETKIMHOGSGMERVEFTYFRVYGFAHGEVLDDEQEFYKESITKLAELISH
RLEEFYTERFGDDVVELIKDSYVPVDSKSLDSQKAYIOLTYVEPYFDEYELKDRVTYFD
RNYGLRTFLFCTPFTPDGRAHGELEPEQHKRKTLLSTDHAFYIKTRIRYCHRETVLT
PVEVALIEDMOKTRELAFATEODPPDAKMLQWLOGSVPTVNOGPLEVAQVFLAEP
EDPKLFRHNMKLRLCERDKCKCEDALRKNKALIGPDQKEYHRELERNYCRDLREALQP
LLTQRLPQLMAPTPPGLRNSLNRASFRAKADL"

BASE COUNT 1273 a 2050 c 1849 g 1200 t
ORIGIN

alignment_scores: length: 2090
Quality: 7136.50

Ratio: 3.998 Gaps: 11
Percent Similarity: 85.407 Percent Identity: 65.311

alignment block:

US-09-737-246-2 x AX173175 ..

Align seg 1/1 to: AX173175 from: 1 to: 6372

```

 2 AlaGluArgArgAlaPheAlaGlnLysIleSerArgThrValAlaAlaG1 18
   :::::::::::::::::::::::::::::|
22 TCCGAGCGCGCGCTTCGCGCACAAAGATCAACAGACGCGTGCGCGCAGA 71
   :::::::::::::::::::::::::::::|
18 uValArgLysGlnIleSerGlyGlnTyrSerGlySerProGlnLeuLeuL 35
   :::::::::::::::::::::::::::::|
72 GGTGCGGAAGCAGAGTGTCCCGGGAACGACAGTGGCTCCCCACTCCAGCA 121
   :::::::::::::::::::::::::::::|
35 ysAsnLeuAsnIleValGlyAsnIleSerHisHisThrThrValProLeu 51
   :::::::::::::::::::::::::::::|
122 GG.....CGCTGCAGCAGCTCCCTGGGGGTCCCACTG 153
   :::::::::::::::::::::::::::::|
52 ThrGluAlaValAspProValAspLeuGluAspTyrLeuIleThrHisPr 68
   :::::::::::::::::::::::::::::|
154 ACTGAAGTTGTCGAGCCCTGGAAGTTGAGGATGTAATTCTGAGCCGGCC 203
   :::::::::::::::::::::::::::::|
68 OleuAlaValAspSerGlyProLeuArgAspLeuIleGluPheProProA 85
   :::::::::::::::::::::::::::::|
204 ACCAGATGCTGAGCCCGCGCCCTCAGGGAGCTGTAGAATTCCAGCTG 253
   :::::::::::::::::::::::::::::|
85 sPAspIleGluValValTyrSerProArgAspCysArgThrLeuValSer 101
   :::::::::::::::::::::::::::::|
254 ATGACTTGGAGCTGCTGCTGCAGACCCCGGGAATGCCGACACGAGCC 303
   :::::::::::::::::::::::::::::|
102 AlaValProGluGluSerGluMetAspProHisValArgAspCysIleAr 118
   :::::::::::::::::::::::::::::|
304 GGGATCCCCAAGGATGAATAACTGGATGCCACAGTGAAGGCGCGGTGA 353
   :::::::::::::::::::::::::::::|
118 gSerTyrThrGluAspTrpAlaIleValIleArgLysTyrHisLysLeuG 135
   :::::::::::::::::::::::::::::|
354 GATGTATATTGAGGACTGGGTCAATTGTCCACAGAAAGTATCAGTACTGA 403
   :::::::::::::::::::::::::::::|
135 lyThrGlyPheAsnProAsnThrLeuAspLysGlnLysGluArgGlnLys 151
   :::::::::::::::::::::::::::::|
404 GTGCAGCATACAGCCCGTCAACACAGACACACGCGGAGCGACAGAGAG 453
   :::::::::::::::::::::::::::::|
152 GlyLeuProLysGlnValPheGluSerAspGluAlaProAspGlyAsnSe 168
   :::::::::::::::::::::::::::::|
454 GGCTTCCCCCGCAGGCTTTTGAGCAGGATGCTTCTTGAGACGAGAGGTC 503
   :::::::::::::::::::::::::::::|
168 rTyrGlnAspAspGlnAspAspLeuLysArgArgSerMetSerIleAspa 185
   :::::::::::::::::::::::::::::|
504 CGGCCCTGAGGACTCGAATGACTCCCGGCGTGGCTCGGCTCCCGGGAAG 553
   :::::::::::::::::::::::::::::|
185 sPThrProArgGlySerTrpAlaCysSerIlePheAspLeuLysAsnSer 201
   :::::::::::::::::::::::::::::|
554 ACACCCCTCGAAGCAGTGGTGCTTACGATCTTCGACCTGAGGAACCTG 603
   :::::::::::::::::::::::::::::|
202 LeuProAspAlaLeuLeuProAsnLeuLeuAspArgThrProAsnGluG1 218
   :::::::::::::::::::::::::::::|
604 GCAGCTGACTCATGTCTGCCCTCTCTGTAGAGCGGGCGGCCCCAGAGA 653
   :::::::::::::::::::::::::::::|
218 uIleAspArgGlnAsnAspAspGlnArgLysSerAsnArgHisLysGluL 235
   :::::::::::::::::::::::::::::|
654 TGTGACCGCGCGCAATGAACCCCTTCGACGCGACACCGCGCCCGGCC 703
   :::::::::::::::::::::::::::::|
235 euPheAlaLeuHisProSerProAspGluGluProIleGluArgLeu 251
   :::::::::::::::::::::::::::::|
704 TGCTCACCTTACCGGACACTGACGAGAGATGAAGCCGTGAACGCTGT 753
   :::::::::::::::::::::::::::::|
252 SerValProAspIleProLysGlnHisPheGlyGlnArgLeuValVally 268
   :::::::::::::::::::::::::::::|
754 AGCGGCCACAGAGCCACCCCGGAGCAGCTTTGGACAAAGGATCTTGGTCAA 803
   :::::::::::::::::::::::::::::|
268 sCysLeuSerLeuLysPheGluIleGluIleGluProIlePheAlaSerL 285
   :::::::::::::::::::::::::::::|
```

```

804 GTGTCTGTGCTCAAGTTCGAGATTGAATTGAGCCCACTTTGGGATCT 853
   ::::::::::::::::::::::::::::::|
285 euAlaLeuTyrAspValLysGlnLysLysLysIleSerGluAsnPheTyr 301
   ::::::::::::::::::::::::::::::|
854 TGGCTGTATGATGTGCGGGAAGAAAGAGATCTCGAGAACTTCTTAC 903
   ::::::::::::::::::::::::::::::|
302 PheAspLeuAsnSerGluGlnMetLysGlyLeuLeuArgProHisValPr 318
   ::::::::::::::::::::::::::::::|
904 TTCGACTGAACCTGGACTCCATGAAGGGGCTGCTTCGGGCTCATGGCAC 953
   ::::::::::::::::::::::::::::::|
318 oProAlaAlaIleThrThrLeuAlaArgSerAlaIlePheSerIleThrT 335
   ::::::::::::::::::::::::::::::|
954 CCACCTTGCCATCTCCACCCCTGGCCGCTCTGCATCTTCTGTGACCT 1003
   ::::::::::::::::::::::::::::::|
335 yrProSerGlnAspValPheLeuValIleLysLeuGluLysValLeuGln 351
   ::::::::::::::::::::::::::::::|
1004 ACCCTCACCTGACATCTCTGCTCATCAAGTTGGAGAAGTGCTTCAG 1053
   ::::::::::::::::::::::::::::::|
352 GlnGlyAspIleGlyGluCysAlaGluProTyrMetIlePheLysGluAl 368
   ::::::::::::::::::::::::::::::|
1054 CAAGGGACATCAGTAGAGTGTGTGAGCCTTACATGGTGTGAAGAAGT 1103
   ::::::::::::::::::::::::::::::|
368 aAspAlaThrLysAsnLysGluLysLeuGluLysLeuLysSerGlnAlaA 385
   ::::::::::::::::::::::::::::::|
1104 GGACACAGCCCAAGAACAAAGAAAGCTAGAGAAGCTGCGCTGGCGCCG 1153
   ::::::::::::::::::::::::::::::|
385 sPcIlnPheCysGlnArgLeuGlyLysTyrArgMetProPheAlaTrpThr 401
   ::::::::::::::::::::::::::::::|
1154 AGCAGTTCTGCACCCGCTGGCGCGCTACCGCATGCCCTTCGCCGTGACG 1203
   ::::::::::::::::::::::::::::::|
402 AlaIleHisLeuMetAsnIleValSerSerAlaGlySerLeuGluArgAs 418
   ::::::::::::::::::::::::::::::|
1204 GCCGTGCACCTTGGCCAACATCGTAGACAGCGCTGGGCACGTGACCCGGA 1253
   ::::::::::::::::::::::::::::::|
418 pSerThrGluValGluIleSerThrGlyLysArgLysGlySerTrpSerG 435
   ::::::::::::::::::::::::::::::|
1254 CTCTGAC.....TCGGAGGGCGAGCGCGGCCCAACCTGGACAG 1291
   ::::::::::::::::::::::::::::::|
435 luArgArgAsnSerSerIleValGlyArgArgSerLeuGluArgThrThr 451
   ::::::::::::::::::::::::::::::|
1292 ACCGCGCGCGT.....CGGGGCCCCAGAGACCGGGCGAGT 1326
   ::::::::::::::::::::::::::::::|
452 SerGlyAspAspAlaCysAsnLeuThrSerPheArgProAlaThrLeuTh 468
   ::::::::::::::::::::::::::::::|
1327 AGTGGGAGCAGACGCTGCACTTCTGTGCTTCCGTCCAGCCACGCTAAC 1376
   ::::::::::::::::::::::::::::::|
468 rValThrAsnPhePheLysGlnGluLysAspArgLeuSerAspGluAspL 485
   ::::::::::::::::::::::::::::::|
1377 TGTCAACAACCTTCTTTAAGCAGGAGGCTGAGCGACTCAGTGACGAGACC 1426
   ::::::::::::::::::::::::::::::|
485 euTyrLysPheLeuAlaAspMetArgArgProSerSerValLeuArgArg 501
   ::::::::::::::::::::::::::::::|
1427 TCTTCAAGTTCCTGGCTGACATGAGGCGCCGTGCTCCCTGTGCGGCGA 1476
   ::::::::::::::::::::::::::::::|
502 LeuArgProIleThrAlaGlnLeuLysIleAspIleSerProAlaProGl 518
   ::::::::::::::::::::::::::::::|
1477 CTACGTCTCTGTACTGCCACGCTCAAGATCGACATTCTCCGGCTCTGA 1526
   ::::::::::::::::::::::::::::::|
518 uAsnProHisTyrCysLeuThrProGluLeuLeuGlnValLysLeuTyrP 535
   ::::::::::::::::::::::::::::::|
1527 AAATCCCCACTTCTGTGCTTCCCTGAGCTGCTTCATATCAAGCCCTTACC 1576
   ::::::::::::::::::::::::::::::|
535 roAspSerArgValArgProThrArgGluIleLeuGluPheProAlaArg 551
   ::::::::::::::::::::::::::::::|
1577 CGGACCCACAGGGCGCGGCCACCAAGAGATTCGAGTTCGCCCGCGCC 1626
   ::::::::::::::::::::::::::::::|
552 AspValTyrValProAsnThrThrTyrArgAsnLeuLeuTyrIleTyrPr 568
   ::::::::::::::::::::::::::::::|
1627 GAAGTCTATAGCCGCCCATACACAGCTACAGGAACCTGCTGACGTGTACCC 1676
   ::::::::::::::::::::::::::::::|
568 oGlnSerLeuAsnPheAlaAsnArgGlnGlySerAlaArgAsnIleThrV 585
   ::::::::::::::::::::::::::::::|
```



```
1185 yrLeuAlaGlyLeuValLeuThrGluLeuAlaValIleLeuAspProAsp 1201
:::|||||:::|||||:::|||||:::|||||:::|||||:::
3335 TCCTAGCTGGCTCTGCTGACGAGCTGCGACTGGCCCTCGAACCTGAG 3384
1202 AlaGluGlyLeuPheGlyLeuHisLysLysValIleAsnMetValHisAs 1218
|||||:::|||||:::|||||:::|||||:::|||||:::
3385 GCTGAAGGGGCAATCCCTGTTGACACAAGAAGGCCATCAGTCTGTGCACAG 3434
1218 nLeuLeuSerSerHisAspSerAspProArgTyrSerAspProGlnIleL 1235
:|||||:::|||||:::|||||:::|||||:::|||||:::
3435 CCTGCTATGTGGCCATGACACTGACCCCGCTACGCCGACACTGTGA 3484
1235 ysAlaArgValAlaMetLeuTyrLeuProLeuIleGlyIleIleMetGlu 1251
|||||:::|||||:::|||||:::|||||:::|||||:::
3485 AGGCTCGTGTGGCCGAGCTGTACTGCCACTGCTTTCGATTGCACGGGAT 3534
1252 ThrValProGlnLeuTyrAspPheThrGluThrHisAsnGlnArgGlyAr 1268
|||:::|||||:::|||||:::|||||:::|||||:::
3535 ACCTTGCCACGGCTGCATGACTTGTGCTGAGGGCCACGTCAGCGGTCAAG 3584
1268 gProIleCysIleAlaThrAspAspTyrGluSerGluSer....Glys 1283
|:::|||||:::|||||:::|||||:::|||||:::
3585 ACTGGCCTCAATGCTTGACTGACACACAGAAGGCGAAGGACATTGCGG 3634
1283 erMetIleSerGlnThrValAlaMetAlaIleAlaGlyThrSerValPro 1299
:::|||||:::|||||:::|||||:::|||||:::
3635 GTACCATCAACCCCTCTGTGGCCATGGCCATGTGTTGGCCCCCTA... 3681
1300 GlnLeuThrArgProGlySerPheLeuLeuThrSer.....Th 1312
|||:::|||||:::|||||:::|||||:::|||||:::
3682 .....GCCCTGGCTCCCGGGCCAGCATCTCCACGAGGGGCCACCAAC 3722
1312 rSerGlyArgGlnHisThrThrPheSerAlaGluSerSerArgSerLeuL 1329
|:::|||||:::|||||:::|||||:::|||||:::
3723 GGCTTCTCGCGCAGGCTGTGCCCTCTCTGCTGAGTCAAGCCGGACCTTGC 3772
1329 euIleCysLeuLeuTyrValLeuLysAsnAlaAspGluThrValLeuGln 1345
|||:::|||||:::|||||:::|||||:::|||||:::
3773 TGGCGTGTGTGTGGGTGTGAAAAAACACGAGCCGGCGCTCTGCAG 3822
1346 LysTrpPheThrAspLeuSerValLeuGlnLeuAsnArgLeuLeuAspLe 1362
:::|||||:::|||||:::|||||:::|||||:::|||||:::
3823 CGCTGGGCCACTGACCTGACACTCCCCACGCTGGGACGCTGTGGACTT 3872
1362 uLeuTyrLeuCysValSerCysPheGluTyrLysGlyLysLysValPheG 1379
|||||:::|||||:::|||||:::|||||:::|||||:::
3873 GCTGTACCTTGGCTGGCTGCTTGTAGTACAAAGGGGAAAAAGGCTTTG 3922
1379 LuArgMetAsnSerLeuThrPheLysSerLysAspMetArgAlaLys 1395
|||||:::|||||:::|||||:::|||||:::|||||:::
3923 AACGCATCAACAGCCTCACATTCAAAAAATCTCTGATATGAAGGCGCGG 3972
1396 LeuGluGluAlaIleLeuGlySerIleGlyAlaArgGlnGluMetValAr 1412
|||||:::|||||:::|||||:::|||||:::|||||:::
3973 CTAGAGGAAGCCATTCTGGGTACCATCGGAGCTCGACAAGAAATGTTGCG 4022
1412 gArgSerArgGlyGlnLeuGluArgSerProSerGlySerAlaPheGlys 1429
|||||:::|||||:::|||||:::|||||:::|||||:::
4023 GCGAAGTGTG.....GAGAGAGCCCG.....TTTGGGA 4051
1429 erGlnGluAsnLeuArgTrpArgLysAspMetThrHisTrpArgGlnAsn 1445
:::|||||:::|||||:::|||||:::|||||:::|||||:::
4052 ATCCAGAGAACGTGCGCTGGCGGAAGAGCGTCACACACTGGAAGCAACC 4101
1446 ThrGluLysLeuAspLysSerArgAlaGluIleGluHisGluAlaLeuIl 1462
:::|||||:::|||||:::|||||:::|||||:::|||||:::
4102 TCAGACCGCGGTGACAAAGACCAAGGATGAATAATGAACACGAGGCGCTTGGT 4151
1462 eaSpGlyAsnLeuAlaThrGluAlaAsnLeuIleIleLeuAspThrLeuG 1479
:::|||||:::|||||:::|||||:::|||||:::|||||:::
4152 GGAAGGGAACCTGGCAACCGAGGCAAGCCTAGTGTCTGTGACACACTGG 4201
1479 IuIleValValGlnThrValSerValThrGluSerLysGluSerIleLeu 1495
|||||:::|||||:::|||||:::|||||:::|||||:::
4202 AGATCATCTGTGCAGACGGTGTATGCTTTCAGAAAGCCCGGAGAGCGTCTTG 4251
1496 GlyGlyValLeuLysValLeuLeuHisSerMetAlaCysAsnGlnSerAl 1512
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::
4252 GGGGCAGTGTGAAGTGTGTGTGTACAGCCTGGGCAGTGCCAGAGTGC 4301
1512 aValTyrLeuGlnHisCysPheAlaThrGlnArgAlaLeuValSerLysP 1529
|:::|||||:::|||||:::|||||:::|||||:::|||||:::
4302 CCTCTCTTGCAGCATGGCCCTGGCCACCCAGAGGGCCCTTGTGTCCAGAT 4351
1529 heProGluLeuLeuPheGluGluGluThrGluGlnCysAlaAspLeuCys 1545
|||||:::|||||:::|||||:::|||||:::|||||:::
4352 TCCCGAGCTGTGTTCGAGAGGACACAGGAGCTGTGTGCCGACTGTGC 4401
1546 LeuArgLeuLeuArgHisCysSerSerSerIleGlyThrIleArgSerHi 1562
|||||:::|||||:::|||||:::|||||:::|||||:::
4402 CTGAGGCTCTACGACACTGTGGCAGCCGCAATCAGCACCATCCGACGCA 4451
1562 sProSerAlaSerLeuTyrLeuLeuMetArgGlnAsnPheGluIleGlyA 1579
|:::|||||:::|||||:::|||||:::|||||:::|||||:::
4452 CGCCAGCGCCTCGCTGACTGCTCATGCGACAGAACTTCAGATCGGCC 4501
1579 snAsnPheAlaArgValLysMetGlnValPrometSerLeuSerSerLeu 1595
:::|||||:::|||||:::|||||:::|||||:::|||||:::
4502 ACAACTTTGCCCGTGTGAAGATGCAAGTCAACCATGTCTCTCGTCCCTG 4551
1596 ValGlyThrSerGlnAsnPheAsnGluGluPheLeuArgArgSerLeuLy 1612
|||||:::|||||:::|||||:::|||||:::|||||:::
4552 GTGGGAGACGACGACAGAACTTCAGTAAGAGCACCTGCGACGTTCACTCAA 4601
1612 sThrIleLeuThrTyrAlaGluGluAspLeuGluLeuArgGluThrThrP 1629
|||||:::|||||:::|||||:::|||||:::|||||:::
4602 AACCATCTTCACCTATGCTGAGGAGAGACATGGGGCTGCGGACAGCACCT 4651
1629 heProAspGlnValGlnAspLeuValPheAsnLeuHisMetIleLeuSer 1645
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::
4652 TCGCAGAGCAGAGTCCAGGACCTGATGTTCAACCTGCACATGATCTGTAGC 4701
1646 AspThrValLysMetLysGluHisGlnGluAspProGluMetLeuIleAs 1662
|||||:::|||||:::|||||:::|||||:::|||||:::
4702 GACACGGTGAAGATGAAGAACACACAGAGGACCTGAGATGCTCATGCA 4751
1662 pleuMetTyrArgIleAlaLysGlyTyrGlnThrSerProGlu..ArgL 1678
|||||:::|||||:::|||||:::|||||:::|||||:::
4752 CCTCATGTACAGAAATGCCCCGGGCTACACAGGCTCAACCGGACCTTGGCC 4801
1678 euThrTyrLeuGlnAsnMetAlaGlyLysHisSerGluArgSerAsnHis 1694
|||||:::|||||:::|||||:::|||||:::|||||:::
4802 TGAACCTGTTGCAGAACATGGCCCGGAAGCACGCGGAGCTGGGCAACCA 4851
1695 AlaGluAlaIleGlnCysLeuValHisSerAlaAlaLeuValAlaGluTyr 1711
|||||:::|||||:::|||||:::|||||:::|||||:::
4852 GCCGAGGCCGCCAGTGCATGTGTACAGCGCGGCCCTCGTGGCTGAGTGA 4901
1711 rLeuSerMetLeuGluAspArgLysTyrLeuProValGlyCysValThrP 1728
|||||:::|||||:::|||||:::|||||:::|||||:::
4902 CCTCGCCCTGCTCGAGAGCAAGCGCACCTGCCGTGGGCTGCGTTTCCCT 4951
1728 heGlnAsnIleSerSerAsnValLeuGluGluSerAlaValSerAspAsp 1744
|||||:::|||||:::|||||:::|||||:::|||||:::
4952 TTCAGAAACATCTCATCCAAAGCTGTAGAGAGATCCGCCATCTCCGAGCAG 5001
1745 ValValSerProAspGluGluGlyIleCysSerGlyLysTyrPheThrGl 1761
:::|||||:::|||||:::|||||:::|||||:::|||||:::
5002 ATCTGTGTCGCCGACGAGGAGGGCTTCTGTCGCGGGAAGCACTTCACATGA 5051
1761 uSerGlyLeuValGlyLeuLeuGluGlnAlaAlaIleSerPheSerMetA 1778
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::
5052 GCTGGGGCTGTGTAGGGTTGCTGGAACAGGCAAGCCGGCTACTTCACCATGG 5101
1778 IaGlyMetTyrGluAlaValAsnGluValTyrLysValLeuIleProIle 1794
:::|||||:::|||||:::|||||:::|||||:::|||||:::

```


5102	GCGGGCTCTACGAGCGCGGTGAATGAGGCTCTACAAGAACCCTCATCCCATC	5151
1795	HISGLUALASNARGASPALALYSLSLEUSERTHRILEHISGLYLSLE ::: :::	1811
5152	CTGGAAGCCACCGCTGACTACAAAGAGCTGGCGCGGTGCACGGCAAACT	5201
1811	UGLNGLUALAPHESERYSLIEVALHISGLINSERTHGLYTRPGLUARGM ::: ::: ::: :::	1828
5202	GCAGGAGCGCTTCACCAAGATCATGCACCAGAGTTCCGGCTGGGAGCGCG	5251
1828	ETPHEGLYTHRTPHREARGVALGLYPHETRYGLYTHRILYSPHLEGLYASP ::: ::: ::: :::	1844
5252	TGTTCCGGGACGATATTCCCGGTGGGCTTCTACGGCGCCCACTTCGGTGAC	5301
1845	LEUASPLUGLNGLUHPHEVALTRYLSGLUPROALALETHRILYSLEUAL ::: ::: ::: :::	1861
5302	CTGGATGAGCAGAGATTGTGTACAAAGAGCCATCGATCAGGAAGCTGGC	5351
1861	AGLUILESERHISARGLEUGLUGLYPHETRYGLYLARGPHEGLYGLUA ::: ::: ::: :::	1878
5352	AGAGATCTCACACCGGCTGGAGAGTTCTACACGAGAGAGATTGGCGACG	5401
1878	SPVALVALGLUVALILELYSASPSEASNPROVALASPLYSCYLSLEU ::: ::: ::: :::	1894
5402	ACGTCGTTGAGATTATCAAGACTCTTACCCTGTGACAACTCCAAAGCTT	5451
1895	ASPROASNLYSALATYRILEGLINILETHRTRYVALGLUPROTPHSPAS ::: ::: ::: :::	1911
5452	GACTCACAAAAGCGCTACATCCAGATCACGSTATGTGAACCGTACTTGA	5501
1911	PTHRTYRGLUMETLYSASPARGILETHRTRYRPHSPASPLYASNTYRASN ::: ::: ::: :::	1928
5502	TACCTACGAGCTCAAGACCGGGTGACCTACTTGACCCGCACATAAGGC	5551
1928	EULARGARGPHEMETTRYRCSYTHRPROPHETHRLEUASPGLYARGALAHIS ::: ::: ::: :::	1944
5552	TTCGCACATCTCTGTCTGCACCGCGTTTCACGCGGATGGCGCGCACAC	5601
1945	GLYGLULEUHHISGLINPHELYSARGLYSTHRILEUTHRTHRSEHNI ::: ::: ::: :::	1961
5602	GGGAGGCTGCCCGAGCAACACAGCGTTAAGACGCTGCTCAGCACCGACCA	5651
1961	SALAPHEPROTYRILEYSTHARGVALASNVALTHRHSIYSGLUGLUT ::: ::: ::: :::	1978
5652	CGCCTTCCCTACATCAAGACTGCGCATCCGTGTGTGCCACCGGAGGAGA	5701
1978	LEILELEUTHRPROLLEGLUVALALALEGLUASPMETGLINYSYSTHR ::: ::: ::: ::: :::	1994
5702	CGGTGCTGACGCCAGTGGAGGTGGCCATCGAGGACATGCAGAAGACACA	5751
1995	GLNGLULEUALAPHALATHRHISGLNASPPROALASPPLYSMETLE ::: ::: ::: ::: :::	2011
5752	CGGAGCTGGCCTTTGCCACCGCAGCAGAACCCACAGATGCTAAGATGCT	5801
2011	UGLNMETVALLEUGLNGLYSERVALGLYTHRTHRVALASNGLNGLYPROL ::: ::: ::: :::	2028
5802	ACAGATGGTGCTTCAGGGCTCTGTAAGGGCCACCGCTGAACCAAGGTCCCC	5851
2028	EUGLUVALALAGLINALPHELUSERGLUTLEPROSERASPPROLYSLEU ::: ::: ::: :::	2044
5852	TGGAGGTGGCCACAGGTGTTTTTACAGAGATCCCGGAAGACCCCAAGCTC	5901
2045	PHEARGHISHISASNLYSLEUARGLEUCYSPHELYSPHETHRYSAR ::: ::: ::: :::	2061
5902	TTCGCGGATCACACAATAATTCGGCTCTGCTCAAGGACTTCTGCAAGAA	5951
2061	GCYSGLUASPALALEUARGLYSASNLYSSERLLEUGLYPROVALGLNL ::: ::: ::: ::: :::	2078
5952	ATGTGAGGATGCGCTGCCGAAAAATTAAGCCCTGATTGGCGCGACACAGA	6001
2078	YSGLUYTRGLNARGGLULEU 2084	
6002	AGGAGTACCAACCGTGAGCTG 6021	

```

seq_name: gb_pr:AB051558

seq_documentation_block:
LOCUS      AB051558                      4311 bp      mRNA      linear      PRI 07-FEB-2001
DEFINITION Homo sapiens mRNA for KIAA1771 protein, partial cds.
ACCESSION  AB051558
VERSION    AB051558.1   GI:12698086
KEYWORDS
SOURCE     Homo sapiens cDNA to mRNA, clone:pj02581.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (sites)
            Nagase,T., Kikuno,R., Hattori,A., Kondo,Y., Okumura,K. and Ohara,O.,
            prediction of the coding sequences of unidentified human genes.
            XIX. The complete sequences of 100 new cDNA clones from brain which
            code for large proteins in vitro
            DNA Res. 7 (6), 347-355 (2000)
JOURNAL    21082932
MEDLINE    2 (bases 1 to 4311)
REFERENCE  Ohara,O., Nagase,T. and Kikuno,R.
AUTHORS    Direct Submission
TITLE       Submitted (22-NOV-2000) Osamu Ohara, Kazusa DNA Research Institute,
            Department of Human Gene Research; 1532-3, Yana, Kisarazu, Chiba
            292-0812, Japan (E-mail:cdnainfo@kazusa.or.jp,
            URL:http://www.kazusa.or.jp/huge, Tel:81-438-52-3913,
            Fax:81-438-52-3914)
FEATURES   Location/Qualifiers
            1..4311
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="pj02581"
                /note="vector:pBluescriptII SK plus"
            1..3910
                /gene="KIAA1771"
                /note="KIAA1771"
                /note="Start codon is not identified. pj02581 cDNA clone
                for KIAA1771 has a 1-bp deletion at the position 498 and a
                110-bp insertion after the position 3119 of the sequence
                of KIAA1771, respectively."
                /codon_start=2
                /product="KIAA1771 protein"
                /protein_id="BAB21862.1"
                /db_xref="GI:12698087"
            /translation="KNEGNHDOGRNSILASYIHYVERLPNTYPNSSPPGGIGGSS
            VHYATMARSAVRPASLNLNRSRSLNSNPDISGTPSPDDEVRSIIGSKGLDRNSW
            NTGPKAAPWGSNDSPSAESTQAMDRCNRMSSHETSSFLQTLTGRLP TKLFEHEL
            ALQWVCGSVRESALQQAQAEFFELMVKSMVHHLFNDKLEAPRKSREPERFMDIAA
            LVSTIASDIVSRFOKDEMERLNTSLAFLNDLLSVMDRGFVSLIKSKYQVSSKIT
            YSLPNPSVLVSLRDLDFLRICSHHYVTNLPCSLTPPASPSVSSATSQSSGFSST
            NVQDQKIANMFELSVPRQOHYLAGLVLTSLAVILDDPAEGLGHLKKVINMWNHLLS
            SHSDPRYSDPQIKARVAMLYPLIGIMETVPOLYDTEETHNORGRPICATDYES
            ESGMSISQTVAMAIAAGTVPQLTRPGSFLITSTSGRQHTTSAESSRLICLVLVLA
            NADETVLQKWFDTLSVLQNLRLIDLILYLCVSCFEYKKGKVEERMSLTFFKSKDMRAA
            LEEAIIIGSIGARQEMVRRSRGQLERSPSGSAFGSQENLRWRKDMTHWRONTEKLDKSH
            AEIEHEALIDGNLATEANLLIIDTLEIVQVQVSYATESKESILGVLKVLHSMACNOG
            AVYLQHCFAFQRAIVSKPEPELLFEETEQAQCLRLRHSSSIGTIRSHASASLYTV
            LMRONFEIGNNFAVKMOVTMSLSSVGTSCDNFNEELRLRSKTLITTYAEDLLELRE
            TEPDOVODIVFNLMHLLSDTVKMEHOEDPEMLIDLMYRIAKGYOTSPDLRTWONNN
            AGKHSERSNHAEAQCLVHASALVAEYLSMLEDRKYLPGCVTFONISSNVLEESAVS
            DDVSPDEEGICSGKYFTESGLGLLEQAAASFMSMGMYEAVNEVYKLIPIHEANRDE
            AKKLSITHGKLQEAFSKIVHQSSTGWERMEFTYFRVGFYGTKFGDLEQEFYKPEAITT
            KLAEISHRLIEGFYGERFEGEDVVEVIKDSNPVDCCKLDPNKAYIQITVEPYFDYEMM
            DRITTFDKNNYLNLRMYCTPFTLDGRAHGLHEQFRKTLITTSAPYIKTRVNVTHR
            KEELIILPIEVAIEDMQKTKQELAFATHODPADPKMLQMWLQGSVGTVNOGPLEVAQ
            VFLSEIPSDPKLFRHNKRLRCFEKDFTRCEDALRKNSLIGPDQKEYQRELERNYHHR
            LKEALQPLINKRKIPQLYKAVLPVTCRDRSFSRMSLRKMDL"
BASE COUNT 1312 a      859 c      920 g      1220 t
ORIGIN

```

alignment_scores:

Quality: 6436.50 Length: 1263
Ratio: 5.121 Gaps: 2
Percent Similarity: 99.525 Percent Identity: 99.446

alignment_block:

US-09-737-246-2 x AB051558 ..

Align seg 1/1 to: AB051558 from: 1 to: 4311

830 LysAsnLeuGluGlyAsnHisAspGlnHisGlyArgAsnSerLeuLeuAl 846
|||||
2 AAAAAGCTGGAGGAATCATGACGACGATGGCAGAAACAGCCTTCTTGC 51
846 aSerTyrIleHisTyrValPheArgLeuProAsnThrTyrProAsnSers 863
|||||
52 ATCATATATTCATATATGTTTCCGCCCTACCAATACTTACCCTAATTCAT 101
863 erSerProGlyProGlyGlyLeuGlyGlySerValHisTyrAlaThrMet 879
|||||
102 CATCACGAGGTCCTGGGGGTTTGGAGGATGATGATGATGATGATGATG 151
880 AlaArgSerAlaValArgProAlaSerLeuAsnLeuAsnArgSerArgSe 896
|||||
152 GCTAGATCTGCGGTGAGACCTGCAAGCCTTAATTAAATCGTTCGGAAG 201
896 rLeuSerAsnSerAsnProAspIleSerGlyThrProThrSerProAspA 913
|||||
202 CCTTAGTATAGCAATCCAGATATATCTGGGACTCCACGTCACACGATG 251
913 spGluValArgSerIleIleGlySerIleGlyLeuAspArgSerAsnSer 929
|||||
252 ATGAAGTTCGATCATCATCGGGAGTAAGGTTTAGATCGCTCCAATTCC 301
930 TrpValAsnThrGlyGlyProLysAlaAlaProTrpGlySerAsnProSe 946
|||||
302 TGGGTTAACACTGGTGGTCCAAAAGCTGCCCATGGGATCCAAACCCAG 351
946 rProSerAlaGluSerThrGlnAlaMetAspArgSerCysAsnArgMets 963
|||||
352 TCCAAGTCAGAAATCAACACAGGCTATGATCGAAGTTGTAATCGTATGT 401
963 erSerHisThrGluThrSerSerPheLeuGlnThrLeuThrGlyArgLeu 979
|||||
402 CTTCGCACACAGACGTCAGATTCTTACAAACATTAAACGGACGCTTA 451
980 ProThrLysLysLeuPheHisGluGluLeuAlaLeuGlnTrpValValCy 996
|||||
452 CCAACTAAAAAGCTTTTTCACGAGAGAGCTGCTTGCAGTGGGTTGTTG 501
996 sSerGlySerValArgGluSerAlaLeuGlnGlnAlaTrpPhePhePheG 1013
|||||
502 CAGTGGCAGCGTTCGGGAATCAGCTTTCACAACAGCCTGGTCTTTTGG 551
1013 IuLeuMetValLysSerMetValHisHisLeuTyrPheAsnAspLysLeu 1029
|||||
552 AATTAAATGGTAAAGACATGGTGCACCACTTATCTTAAATGATAAACTT 601
1030 GluAlaProArgLysSerArgPheProGluArgPheMetAspPheLeuAl 1046
|||||
602 GAGGCTCCAAAGAAAGTCGTTTTCAGAACGTTTCATGATGACATTGC 651
1046 aAlaLeuValSerThrIleAlaSerAspIleValSerArgPheGlnLysA 1063
|||||
652 AGCTCTGTGACAGCATGCTAGTAGATATAGTTTCACGATTTCCAGAAG 701
1063 sPThrGluMetValGluArgLeuAsnThrSerLeuAlaPhePheLeuAsn 1079
|||||
702 ACACAGAAATGCTTGAGAGACTCAATACAAAGCCTTGCAATCTTTCTCAAT 751
1080 AspLeuLeuSerValMetAspArgGlyPheValPheSerLeuIleLysSe 1096
|||||
752 GATCTGTGTCCTGTTATGGACAGAGGATTTGTTTACCTTATAAAGTC 801

1096 rCysTyrLysGlnValSerSerLysLeuTyrSerLeuProAsnProSerV 1113
|||||
802 CTGCTATAAACAGGTGCTTCAAGCCTTTACTCATTTACCGAATCCAGATG 851
1113 alLeuValSerLeuArgLeuAspPheLeuArgIleIleCysSerHisGlu 1129
|||||
852 TTCTGGTGTCCCTTGAGGCTGGATTTTCTACGAATCATCTGCAGTCATGAG 901
1130 HisTyrValThrLeuAsnLeuProCysSerLeuLeuThrProProAlaSe 1146
|||||
902 CACTATGTTACATTAACCTTACCCCTGCAGCTTACTTACTCCACCTGCATC 951
1146 rProSerProSerValSerSerAlaThrSerGlnSerSerGlyPheSerT 1163
|||||
952 TCCATCACCTTCTGTCTCTGCAACATCTCAGAGTTCTGGATTTTCTA 1001
1163 hrAsnValGlnAspGlnLysIleAlaAsnMetPheGluLeuSerValPro 1179
|||||
1002 CGAATGTACAAGACCAAAAGATTGCCAAATATGTTGAATTATCCGTGCT 1051
1180 PheArgGlnGlnHisTyrLeuAlaGlyLeuValLeuThrGluLeuAlaVal 1196
|||||
1052 TTCGCCCAACAGCATTTATTTGGCAGAGACTTGTGTTAAACAGAGCTGGCTGT 1101
1196 IleLeuAspProAspAlaGluGlyLeuPheGlyLeuHisLysLysValI 1213
|||||
1102 CATTTTAGACCCCTGATGCTGAAGGACTGTTTGAATTGCATAAGAAAGTCA 1151
1213 leAsnMetValHisAsnLeuLeuSerSerHisAspSerAspProArgTyr 1229
|||||
1152 TCATATATGGTACACAATTTACTCTCCAGTCACGACTCAGACCCGCGGTAC 1201
1230 SerAspProGlnIleLysAlaArgValAlaMetLeuTyrLeuProLeuI 1246
|||||
1202 TCTGACCCCTCAGATTAAGGCTCGAGTGGCCATGTTGATCTACCTCTGAT 1251
1246 eGlyIleIleMetGluThrValProGlnLeuTyrAspPheThrGluThrH 1263
|||||
1252 TGGTATTTATCATGGAACCTGTACCTCAGCTGTATGATTTTACAGAAACTC 1301
1263 isAsnGlnArgGlyArgProIleCysIleAlaThrAspAspTyrGluSer 1279
|||||
1302 ACAATCAACGAGGAAGACCAATTTGTATAGCCACTGATGATTTATGAAGT 1351
1280 GluSerGlySerMetIleSerGlnThrValAlaMetAlaIleAlaGlyTh 1296
|||||
1352 GAGAGCGGAAGTATGATAGACCCAGACCGTGGCCATGGCAATCGCAGGAGC 1401
1296 rSerValProGlnLeuThrArgProGlySerPheLeuLeuThrSerThrs 1313
|||||
1402 ATCGGTCCCTCAACTAACAAAGCCCTGGCAGTTCTCCTCACGTCACAGCA 1451
1313 erGlyArgGlnHisThrThrPheSerAlaGluSerSerArgSerLeuLeu 1329
|||||
1452 GTGGCAGGCAACACACTACTTTTTCAGCAGATCAAGTCGAAGCCTTTTG 1501
1330 IleCysLeuLeuThrValLeuLysAsnAlaAspGluThrValLeuGlnLy 1346
|||||
1502 ATCTGTCTACTTTGGGTTCTCAAAAATGCAGATGAACAAGTTCTACAGAA 1551
1346 sTrpPheThrAspLeuSerValLeuGlnLeuAsnArgLeuLeuAspLeuL 1363
|||||
1552 GTGGTTTACAGATCTCTCAGTCTTGCAGCTAAACCGGCTATTAGATCTGC 1601
1363 eUtyrLeuCysValSerCysPheGluTyrLysGlyLysLysValPheGlu 1379
|||||
1602 TTTATCTCTGTGTCTGCTTTGAGTATAAAGGAAAAAGTGTGTA 1651
1380 ArgMetAsnSerLeuThrPheLysLysSerLysAspMetArgAlaLysLe 1396
|||||
1652 CGAATGAATAGCTTGACCTTTAAGAAATCAAAAGACATGAGAGCAAGCT 1701

1396 uGIuGluaIaIleuGlSerIleGIyAlaArgGInglumetValArgA 1413
|||||
1702 TGAAGAAGCTATTCTTGGAGCATAGGTGCCAGCAAGAAATGTTACGGC 1751
1413 rGSerArGlyGInleuGlUArgSerProSerGlySerAlaPheGlySer 1429
|||||
1752 GAAGCCGAGGACAGCTCGAGAGAAGCCCATCTGGAAGTGCCCTTGGAAGT 1801
1430 GInglUAsnleuArGTrPArgLySAspMetThrHsTrPARgInAsnTh 1446
|||||
1802 CAAGAAAAATTGAGGTGGAGGAAAGATATGACTCACTGGCGTCAAAACAC 1851
1446 rGIuLySleuAsPlySserArgAlaGluiIeGIuHISgluaIaIleuIleA 1463
|||||
1852 AGAGAAGCTTGACAAATCAAGACAGAGATTTGAACACGAAGCACTGATTG 1901
1463 sPGlyAsnleuAlaThrGluAlaAsnleuIleIleleuAspThrleuGlU 1479
|||||
1902 ATGGAACCTGGCTACAGAACCAACCTAATCATTTTATACATTAGAG 1951
1480 IleValValGInThrValSerValThrGlUserLySglUserIleleuGI 1496
|||||
1952 ATTGTTGTTcAGACCGTTCTGTGAACGGAATCCAAGAGAGCATTTCTTG 2001
1496 yGIyValleuLySValleuUeuhISserMetAlaCysAsnGInserAlaV 1513
|||||
2002 TGGAGTGTAAAAAGTGTACTACACAGCATGGCTGTAAACCAAGTGcAG 2051
1513 aITyrlleuGInHIScysPheAlaThrGInArgAlaIleuValSerLySphe 1529
|||||
2052 TTATCTACACACTGTTTGCTACACAGAGAGCCTTGGTTCAAGTTT 2101
1530 ProGIuLeuLeuPhegluGlUglUThrGlUgInCysAlaAspleuCysLe 1546
|||||
2102 CCTGAACCTCTTATTGAAAGACAGACAGAGCAGTGCTGATTATGcCT 2151
1546 uArgleuLeuArGhIScysSerSerSerIleGIyThrIleArgSerHISp 1563
|||||
2152 CAGGCTTCTCCGACACTGTAGCAGTAGCATCGGTACAAATACGGGTcACACG 2201
1563 roSerAlaSerleuTyrlleuLeuMetArgGInAsnPhegluiIeGIyAsn 1579
|||||
2202 CCAGTGCCTCCCTTTACCTACTAATGAGGCAAACTTTGAGATTGGGAAT 2251
1580 AsnPheAlaArgValLySmetGInValPrometSerleuSerSerleuVa 1596
|||||
2252 AACTTTGCCAGGTTAAAAATGCAGGTAAACAATGTCACTATCCTCTTGGT 2301
1596 IGIyThrSerGInAsnPheAsnGIuGlUPheleuArgArgSerleuLysT 1613
|||||
2302 GGGCACATCTCAGAATTTTAAAGAAGATTTCTTAAGACCTTCTCTAAAGA 2351
1613 hrIleleuThrTyrlAlaGlUAspleuGIuLeuArgGIuThrThrphe 1629
|||||
2352 CTATATTGACATATGTCTGAAGAGATCTGGAATTGAGGGAACAACATTTT 2401
1630 ProAspGInValGInAspleuValPheAsnleuHISmetIleleuSerAs 1646
|||||
2402 CCTGATCAGGTCAGAGATCTGGTTTCAATCTCCATATGATTCTTTCTGA 2451
1646 pThrValLySmetLySglUHISglInglUAspProGIumetleuIleAspL 1663
|||||
2452 TACTTGAAAAATGAAGAACACAGGAGGATCCTGAATGTGATTGATC 2501
1663 eumetTyArgIleAlaLySglTyrgInThrSerProGIu...Argleu 1678
|||||
2502 TAATGTACAGAAATTGCCAAGGGTTACAGACCTCTCCAGATCTGCGATTG 2551
1679 ThrTripleuGInAsnMetAlaGIyLySHISserGIuArgSerAsnHISAl 1695
|||||
2552 ACCTGGTGCAGACATGGCAGGCAAGCACTCAGAACGAAGCAATCATGC 2601
1695 aGIuAlaIaIaGInCysleuValHISserAlaAlaIeUValAlaGIuTyrl 1712

|||||
2602 TGAAGCTGCACAGTGTCTAGTCCACTCAGACGACACTGTGTGCTGAATATT 2651
1712 euSerMetleuGIuAspArgLySTyrlleuProValGIyCysValThrPhe 1728
|||||
2652 TGAGCATGCTGGAGGACCGGAATATCTTCTGCGGATGTGTAACATTTT 2701
1729 GInAsnIleSerSerAsnValleuGIuGlUserAlaValSerAspAspVa 1745
|||||
2702 CAGAAATATTTCACTTAATGTTTGAAGAAGATCTGCGGTCTCAGATGATGT 2751
1745 IValSerProAspGIuGIuGIyIleCysSerGIyLySTyrlPheThrGIuS 1762
|||||
2752 GGTATCTCCAGATGAAGAAGGTATCTGCTCTGGAATAATACTTTACTGAGT 2801
1762 eRGIyLeuValGIyLeuLeuGIuGInAlaAlaAlaSerPheSerMetAla 1778
|||||
2802 CAGACTTGTTGGGATTACTGGAACAAGCAAGCTGCTTCCTTCTATGGCT 2851
1779 GIyMetTyrgIuAlaValaIaSnGIuValTyrlLySValleuIleProIleH 1795
|||||
2852 GGCATGTATGAAGCAGTTAATGAAGTTTACAAGTACTTATTCCTATTCA 2901
1795 sGIuAlaAsnArgAspAlaLySlySleuSerThrIleHISGIyLySleuG 1812
|||||
2902 TGAAGCTAATCGGGATGCAAAGAACTATCCACAATTCAATGTAAACTTC 2951
1812 InGIuAlaPheSerLySIleValHISGInSerThrGIyTrPGIuArgMet 1828
|||||
2952 AAGAAGCATTCAGCAAAATGTTTCATCAGAGTACTGGCTGGGAGCGGATG 3001
1829 PheGIyThrTyrlPheArgValGIyPheTyrgLyThrLySPheGIyAsple 1845
|||||
3002 TTTGGACACCTATTTCGTGTGTTTATGGAACCAAGTTGCGGGATTT 3051
1845 uAspGIuGInglUPheValTyrlLySglUProAlaIleThrLySleuAlaG 1862
|||||
3052 GGATGAACAGAAATTTGTTTACAGAGAGCCTGCAATTAACCAACTTGCAG 3101
1862 IuIleSerHISArgleuGIuGIyPheTyrgLyGIuArgPheGIyGIuAsp 1878
|||||
3102 AGATATCTCACAGATTTGAGGAGATTTACGGAGAAAGATTGGAGAGAGAT 3151
1879 ValValGIuValIleLySAspSerAsnProValAsPlyScysLySleuAs 1895
|||||
3152 GTGGTGAAGTAATCAAAAGACTTAATCCTGTAGACAAAGTGTAAATTAGA 3201
1895 pProAsnLySAlaTyrlIeGIuIleThrTyrlValGIuProTyrlPheAspT 1912
|||||
3202 TCCTAACAGGcATATATTTCAGATTACTATGTGAGCCATTACTTTGACA 3251
1912 hrTyrgIuMetLySAspArgIleThrTyrlPheAsPlySAsnTyrlAsnleu 1928
|||||
3252 CATATGAGATGAAGACAGAAATCACCTAATTTCGACAAAAAATTACAATCTT 3301
1929 ArgArgPheMetTyrcysThrProPheThrleuAspGIyArgAlaHISGI 1945
|||||
3302 CGTGCATTCACTACTGTACACCCCTTACTTTAGATGGCCGTGCCATGG 3351
1945 yGIuLeuHISglInPheLySArgLySThrIleleuThrThrSerHISa 1962
|||||
3352 GGAACCTCATGACAATTCAAAAGGAAGACCATTTCTGACTACGTCTCATG 3401
1962 IAPheProTyrlIleLySThrArgValaIaSnValThrHISLySglUgluIle 1978
|||||
3402 CCTTTCCTTATATTAAACAAGGGTCAATGTCACTCATTAAGACAGATC 3451
1979 IleleuThrProIleGIuValaIaIeGIuAspMetGInLySlySThrGI 1995
|||||
3452 ATCTTAACACCAATTGAAGTTGCTATTGAGGACATGCAGAAAAAGACACA 3501
1995 nGIuLeuAlaPheAlaThrHISGInAspProAlaAspProLySmetleuG 2012
|||||

3502 GGAGTTGGCATTGTGCAACACATCAGGATCCCGCAGACCCCAAAATGCTTC 3551
2012 lmetValleuGInglySerValGlyThrThrValAsnGInglyProleu 2028
|||||
3552 AGATGGTACTCCAGGACTCTGTAGCACACAGTGAATCAGGGGCTTTG 3601
2029 GluValAlaGlnValPheLeuSerGluIleProSerAspProlysLeuPh 2045
|||||
3602 GAAGTTGCCAGGTTTCTGTCTGGAATACCTAGTGACCCAAAGCTCTT 3651
2045 eargHisHisAsnLysLeuArgLeuCysPheLysAspPheThrLysArgC 2062
|||||
3652 CAGACATCATATAATACTGCGACTCTGCTTTAAAGATTTTACTAAAAAGT 3701
2062 ysgLAspAlaLeuArgLysAsnLysSerLeuIleGlyProValGlnLys 2078
|||||
3702 GTGAAGATGCCCTTAAGAAAAAATAAAGACTTAATTGGCCGGATCAAAAG 3751
2079 GluTyrGlnArgGluLeu..GlyLysLeuSerSerPro 2090
|||||
3752 GAGTATCAAAAGGAACTGAGAGAAAAACTATCATCGCCT 3789

seq_name: gb_pat:AX172980

seq_documentation_block:

LOCUS AX172980 4200 bp DNA linear PAT 03-JUL-2001
DEFINITION Sequence 107 from Patent WO0142297.
ACCESSION AX172980
VERSION AX172980.1 GI:14597958
KEYWORDS
SOURCE .
ORGANISM synthetic construct.
artificial construct.
REFERENCE 1 (bases 1 to 4200)
AUTHORS Lu,P., Garman,J.D. and Candia,A.F.
TITLE Clasp-3 transmembrane protein
JOURNAL Patent: WO 0142297-A 107 14-JUN-2001;
Arbor Vita Corporation (US)
FEATURES
location/Qualifiers
source 1..4200
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="polynucleotide fragment"

BASE COUNT 1262 a 844 c 907 g 1187 t
ORIGIN

alignment_scores:

Quality: 6171.50 Length: 1226
Ratio: 5.088 Gaps: 2
Percent Similarity: 98.940 Percent Identity: 98.206

alignment_block:

us-09-737-246-2 x AX172980 ..

Align seg 1/1 to: AX172980 from: 1 to: 4200

866 GlyProGlyGlyLeuGlyGlySerValHisTyrAlaThrMetaIaargSe 882
|||||
3 GGGCCCGGTCCCGCTCGGTGCAGCGGTATCGATAAGCTTGATATCGAATTC 52
882 rAlaValArgProAlaSerLeuAsnLeuAsnArgSerArgSerLeuSera 899
|||||
53 GGCA.....CGAGTAAATCGTCTCGAAGCCTTAGTA 84
899 snSerAsnProAspIleSerGlyThrProThrSerProAspAspGluVal 915
|||||
85 ATAGCAATCCAGATATATCTGGGACTCCACGTCACACAGATGATGAAGTT 134
916 ArgSerIleIleGlySerLysGlyLeuAspArgSerAsnSerTrpValas 932
|||||
135 CGATCAATCATCGGAGTAGAGGTTAGATCGCTCCAATTCCTGGGTTAA 184

932 nThrGlyGlyProLysAlaAlaProTrpGlySerAsnProSerProsera 949
|||||
185 CACTGGTGCTCCAAAGCTGCCCATATGGGATCCAAACCCAGTCCAAAGTG 234
949 lAgIuSerThrGlnAlaMetAspArgSerCysAsnArgMetSerSerHis 965
|||||
235 CAGAAATCAACACAGGCTATGATCGAAGTTGATATCGTATGCTTCGCAC 284
966 ThrGluThrSerSerPheLeuGlnThrLeuThrGlyArgLeuProThrLy 982
|||||
285 ACAGAGACGTCAGGTTCTTACAAACATTAAACGGGAGCGCTTACCAACTAA 334
982 slLysLeuPheHisGlnGluLeuAlaLeuGlnTrpValValCysSerGlyS 999
|||||
335 AAAGCTTTTTCACGAGGAGCTGGCTTTCAGTGGCTGTTTGCAGTGGCA 384
999 erValArgGluSerAlaLeuGlnGlnAlaTrpPhePhePheGluLeuMet 1015
|||||
385 GCGTTCCGGAATCAGCTTTGCCAACACAGCCCTGTTCTTTTGAATTAATG 434
1016 ValLysSerMetValHisHisLeuTyrPheAsnAspLysLeuGluAlaPr 1032
|||||
435 GTAAGAGACATGTGTCACCATTTATCTTAAATGATAAAGTTGAGGCTCC 484
1032 cArgLysSerArgPheProGluArgPheMetAspAspIleAlaAlaLeuV 1049
|||||
485 AAGGAAAAAGTCGTTTCCAGAACGTTTCATGATGATGACATTGCAGCTCTTG 534
1049 alSerThrIleAlaSerAspIleValSerArgPheGlnLysAspThrGlu 1065
|||||
535 TCAGCACGATTGCTAGTATATAGTTTCACGATTTTCAGAAAGACACAGAA 584
1066 MetValGluArgLeuAsnThrSerLeuAlaPhePheLeuAsnAspLeuE 1082
|||||
585 ATGGTTGAGAGACTCAATACAAAGCCTTGCAATCTCTTCAATGATCTGTT 634
1082 uSerValMetAspArgGlyPheValPheSerLeuIleLysSerCysTyrL 1099
|||||
635 GTCTGTTATGACAGAGAGATTGTTTATAGCCTTAATAAGTCCCTGCTATA 684
1099 ysgLInValSerSerLysLeuTyrSerLeuProAsnProSerValLeuVal 1115
|||||
685 AACAGGTGCTTCAAGCTTTACTCATTTACCGAATCCAGTGTTCTGCTG 734
1116 SerLeuArgLeuAspPheLeuArgIleIleCysSerHisGlnHisTyrVa 1132
|||||
735 TCCTTGAGGCTGGAATTTTCTACGAATCATCTGCAGTCATGAGCAGTATGT 784
1132 lThrLeuAsnLeuProCysSerLeuLeuThrProProAlaSerProSerP 1149
|||||
785 TACATTTAACTTACCCTGCAGCTTACTTACCTCCAGCTGCATCTCCATCAC 834
1149 roSerValSerSerAlaThrSerGlnSerSerGlyPheSerThrAsnVal 1165
|||||
835 CTTCGTGTTCTTCTGCAACATCTCAGAGTTCTGCAATTTCTACGAATGTA 884
1166 GluAspGlnLysIleAlaAsnMetPheGluLeuSerValProPheArgL 1182
|||||
885 CAAGACCAAAAGATGCAATATATGTTGAATTAATCCGCTTCCGCCA 934
1182 ngLInHisTyrLeuAlaGlyLeuValLeuThrGluLeuAlaValIleLeuA 1199
|||||
935 ACAGCATTAATTTGGCAGGACTTGTGTTAACAGAGCTGGCTGTCAATTTAG 984
1199 spProAspAlaGluGlyLeuPheGlyLeuHisLysLysValIleAsnMet 1215
|||||
985 ACCCTGATGCTGAAGGACTGTTTGGATTCATTAAGAAAGTCATCAATATG 1034
1216 ValHisAsnLeuLeuSerSerHisAspSerAspProArgTyrSerAspPr 1232
|||||
1035 GTACACAATTTACTCTCCAGTCACGACTCAGACCCGCGGTACTGACCC 1084
1232 ogLInIleLysAlaArgValAlaMetLeuTyrLeuProLeuIleGlyIleI 1249

|||||
1085 TCAGATAAAGCTCGAGTGGCCATGTGTATCTACCTCTGATTTGATTTA 1134
1249 IemetGluThrValProGlnLeuTyrAspPheThrGluThrHisAsnGln 1265
|||||
1135 TCATGGAAACTGTACCTCAGCTGTATGATTTTACAGAAACTCACAAATCAA 1184
1266 ArgGlyArgProIleCysIleAlaThrAspAspTyrGluSerGluSerGln 1282
|||||
1185 CGAGGAAGACCAATTGTATAGCCACTGATGATTTATGAAGTGAGAGCGG 1234
1282 ySerMetIleSerGlnThrValAlaMetAlaIleAlaGlyThrSerValP 1299
|||||
1235 AAGTATGATAGCCAGACCGTTGCCATGGCAATCGAGGACATCGGTCC 1284
1299 roGlnLeuThrArgProGlySerPheLeuLeuThrSerThrSerGlyArg 1315
|||||
1285 CTCAACTTAACAAGCGCTGGCAGTTTCCCTCCTCAGCTCAACGAGTGGCAGG 1334
1316 GlnHisThrThrPheSerAlaGluSerSerArgSerLeuLeuIleCysIle 1332
|||||
1335 CAACACACTACCTTTTCAGCAGAAATCAAGTCGAGCCTTTGATCTGTCT 1384
1332 uLeuTrpValLeuLysAsnAlaAspGluThrValLeuGlnLysTrpPheT 1349
|||||
1385 ACTTTGGGTTCTCAAAATGCAGATGAACACAGTTCTACAGAAGTGGTTTA 1434
1349 hrasPLeuSerValLeuGlnLeuAsnArgLeuLeuAspLeuLeuTyrLeu 1365
|||||
1435 CAGATCTCTCAGTCTTGACAGCTAAACCGGCTATTAGATCTGCTTTATCTC 1484
1366 CysValSerCysPheGluTyrLysGlyLysLysValPheGluArgMetAs 1382
|||||
1485 TGTGTGTCCTTGCTTGAGTATAAGGGAAAAAGTGTTTGAACGAATGAA 1534
1382 nSerLeuThrPheLysLysSerLysAspMetArgAlaLysLeuGluGluA 1399
|||||
1535 TAGCTTGACCTTTAAGAAATCAAAAAGACATGAGAGCAAAAGCTTGAAGAAG 1584
1399 IaIleLeuGlySerIleGlyAlaArgGlnGluMetValArgArgSerArg 1415
|||||
1585 CTATTCTTGGGAGCATAGTGGCCAGGCAAGAAATGTTACGGCGAAGCCGA 1634
1416 GlyGlnLeuGluArgSerProSerGlySerAlaPheGlySerGlnGluAs 1432
|||||
1635 GGACAGCTCGAGAGAGCCCATCTGGAAGTGCCTTTGGAAGTCAAGAAAA 1684
1432 nLeuArgTrpArgLysAspMetThrHisTrpArgGlnAsnThrGluLysL 1449
|||||
1685 TTTGAGGTGGAGGAAGATATGACTCAGCTGGCGTCAAAAACACAGAGAAGC 1734
1449 euAspLysSerArgAlaGluIleGlnHisGluAlaLeuIleAspGlyAsn 1465
|||||
1735 TTGACAAATCAAGAGCAGAGATTGAACACGAAGCAGTGAATTGATGAAAC 1784
1466 LeuAlaThrGluAlaAsnLeuIleIleLeuAspThrLeuGluIleValVa 1482
|||||
1785 CTGGCTACAGAAAGCAAACTAATCATTTTAGATACATTAGAGATTGTGT 1834
1482 IGI nThrValSerValThrGluSerLysGluSerIleLeuGlyGlyValL 1499
|||||
1835 TCAGACCGTTTCTGTACGGGATCCAAAGAGAGCATTTCTGGTGAGGTGC 1884
1499 euLysValLeuLeuHisSerMetAlaCysAsnGlnSerAlaValTyrLeu 1515
|||||
1885 TAAAGTGTCTACTACACAGCATGGCCTGTAAACCAAAGTGCAGTTATCTTA 1934
1516 GlnHisCysPheAlaThrGlnArgAlaLeuValSerLysPheProGluLe 1532
|||||
1935 CAACACTGTTTGTCTACACAGAGAGCCTTGGTTTCAAAAGTTTCCCTGAAC 1984
1532 uLeuPheGluGluGluThrGluGlnCysAlaAspLeuCysLeuArgLeuL 1549
|||||

1985 CTTATTGGAAGAGAGACAGACAGTGTGCTGATTTATGCCCTCAGGCTTC 2034
1549 euArgHisCysSerSerSerIleGlyThrIleArgSerHisProSerAla 1565
|||||
2035 TCCGACACTGTAGCAGTAGCATGGTACAAATACGCTCACACCCAGTGC 2084
1566 SerLeuTyrLeuLeuMetArgGlnAsnPheGluIleGlyAsnAsnPheAl 1582
|||||
2085 TCCCTTAACCTACTAATGAGGCAAAACTTTGAGATTTGGGAATAACTTTGC 2134
1582 aArgValLysMetGlnValPrometSerLeuSerSerLeuValGlyThrs 1599
|||||
2135 CAGGTTAAATGACAGGTACCAATGTCACATCTCCTCTTGGTGGCACAT 2184
1599 ergLAsnPheAsnGluGluPheLeuArgArgSerLeuLysThrIleLeu 1615
|||||
2185 CTCAGAAATTTAATGAAGAATTTCTTAAGACGTTCTCTAAGACTATATTG 2234
1616 ThrTyrAlaGluGluAspLeuGluLeuArgGluThrThrPheProAspGl 1632
|||||
2235 ACATATGCTGAAGAAGATCTGGAATTTGAGGGAACAACATTTCTCGATCA 2284
1632 nValGlnAspLeuValPheAsnLeuHisMetIleLeuSerAspThrValL 1649
|||||
2285 GGTCCAGATCTGTGTTTCAATCTCCATATGATTTCTTCTGATACTGTGA 2334
1649 ySerMetLysGluHisGlnGluAspProGluMetLeuIleAspLeuMetTyr 1665
|||||
2335 AAATGAAGAACACACAGAGAGATCTGAAATGTGATGTGATCTAATGTAC 2384
1666 ArgIleAlaLysGlyTyrGlnThrSerProGlu..ArgLeuThrTriple 1681
|||||
2385 AGAATTTGCCAAGGGTTTACCAGACCTCTCCAGATCTGCGATTTGACCTGTT 2434
1681 uGlnAsnMetAlaGlyLysHisSerGluArgSerAsnHisAlaGluAlaA 1698
|||||
2435 GCAGAACATGGCAGGCAAGCAGCTCAGAACGAACAAATCATGCTGAAGCTG 2484
1698 IaGlnCysLeuValHisSerAlaAlaLeuValAlaGluTyrLeuSerMet 1714
|||||
2485 CACAGTGTCTAGTCCACTCAGCAGCACTTGTGCTGAATATTGAGCATG 2534
1715 LeuGluAspArgLysTyrLeuProValGlyCysValThrPheGlnAsnIl 1731
|||||
2535 CTGAGAGACCGAAATATCTTCTGTGGGATGTGTAACAATTTCAGAATAT 2584
1731 eSerSerAsnValLeuGluGluSerAlaValSerAspAspValValSerP 1748
|||||
2585 TTCATCTAATGTTTTAGAAGAATCTGCCGCTCTCAGATGATGTGTATCTC 2634
1748 roAspGluGluGlyIleCysSerGlyLysTyrPheThrGluSerGlyLeu 1764
|||||
2635 CAGATGAAGAGGTATCTGCTCTGGAATAATACTTTACTGAGTCAGGACTT 2684
1765 ValGlyLeuLeuGluGlnAlaAlaAlaSerPheSerMetAlaGlyMetTy 1781
|||||
2685 GTGGATTACTGGAACAAGCAGCTGCTTCTCTATAGCTGGCATGTA 2734
1781 rGluAlaValAsnGluValTyrLysValLeuIleProIleHisGluAlaA 1798
|||||
2735 TGAAGCAGTTAATGAAGTTTACAAGTACTTATTTCTATTCATGAAGCTA 2784
1798 snArgAspAlaLysLysLeuSerThrIleHisGlyLysLeuGlnGluAla 1814
|||||
2785 ATCGGAGATGCAAGAAACTATCCACAATTCATGTGTAACCTCAAGAAGCA 2834
1815 PheSerLysIleValHisGlnSerThrGlyTrpGluArgMetPheGlyTh 1831
|||||
2835 TTCAGCAAAATTTGTCATCAGAGTACTGCGCTGGAGCGGATTTTGGCAC 2884
1831 rTyrPheArgValGlyPheTyrGlyThrLysPheGlyAspLeuAspGluG 1848
|||||
2885 CTATTTTCGTGTGTTTATGGAACCAAGTTGCGGGATTTGGATGAAC 2934

1848 lngluphevalTyrLysGluProAlaIleThrLysLeuAlaGluIleSer 1864
|||||
2935 AAGAATTGTTACAGAGAGCCTGCATATACCAACTTGCAGAGATATCT 2984
1865 HisArgLeuGluGlyPheTyrGlyGluArgPheGlyGluAspValValGI 1881
|||||
2985 CACAGATTGGAGGGATTTTACGGAGAAAGATTGGAGAGGATGTGGTTGA 3034
1881 uValIleLysAspSerAsnProValAspLysCysLysLeuAspProAsnL 1898
|||||
3035 AGTAATCAAGACTCTAATCCTGTAGACAAAGTGTAAATTAGATCTTAACA 3084
1898 ysAlaTyrIleGlnIleThrTyrValGluProTyrPheAspThrTyrGlu 1914
|||||
3085 AGGATATATTCAGATTACCTATGTGAGCCATACTTGCACACATATGAG 3134
1915 MetLysAspArgIleThrTyrPheAspLysAsnTyrAsnLeuArgArgPh 1931
|||||
3135 ATGAAGACAGAAATCACCTATTTCGACAAAAATTACATCTTCGTGATTT 3184
1931 eMetTyrCysThrProPheThrLeuAspGlyArgAlaHisGlyGluLeuH 1948
|||||
3185 CATGACTGTACACCCCTTACTTAGATGGCCGTGCCCATGGGGAATTC 3234
1948 isGluGlnPheLysArgLysThrIleLeuThrThrSerHisAlaPhePro 1964
|||||
3235 ATGAACAATTCAAAAGGAAGACCATTTCTGACTACGCTCATGCTTTCT 3284
1965 TyrIleLysThrArgValAsnValThrHisLysGluGluIleLeuTh 1981
|||||
3285 TATATTAAACAAGGGTCAATGTCACTCATTAAGAAGAGATCATCTTAAC 3334
1981 rProIleGluValAlaIleGluAspMetGlnLysLysThrGlnGluLeuA 1998
|||||
3335 ACCAATTGAAGTTGCTATTGAGACATGCAGAAAAAGACACAGAGTTGG 3384
1998 lApheAlaThrHisGlnAspProAlaAspProLysMetLeuGlnMetVal 2014
|||||
3385 CATTTGCAACACATCAGATCCCGCAGACCCCAAAATGCTTCAGATGGTA 3434
2015 LeuGlnGlySerValGlyThrThrValAsnGlnGlyProLeuGluValAl 2031
|||||
3435 CTCACAGGATCTGTAGGCACACAGTGAATCAGGGGCTTTGGAAAGTTGC 3484
2031 aGlnValPheLeuSerGluIleProSerAspProLysLeuPheArgHisH 2048
|||||
3485 CCAGGTTTCTGTCTGAAATACCTAGTGACCCCAAGCTCTTCAGACATC 3534
2048 tAsnLysLeuArgLeuCysPheLysAspPheThrLysArgCysGluAsp 2064
|||||
3535 ATATATAACTGCGACTGTGCTTAAGATTTTACTAAAAGGTGGAAGAT 3584
2065 AlAlaLeuArgLysAsnLysSerLeuIleGlyProValGlnLysGluTyrGI 2081
|||||
3585 GCCTTAAGAAAAAATAAGAGCTTAATTGGGCCGTTCAAAAGGAGTATCA 3634
2081 nArgGluLeuGlyLysLeuSerSerPro 2090
|||||
3635 AAGGGAATGGGGAACATATCTTGCCT 3662

seq_name: gb_pat:AX172880

seq_documentation_block: 4144 bp DNA linear PAT 03-JUL-2001
LOCUS AX172880
DEFINITION Sequence 7 from Patent WO0142297.
ACCESSION AX172880
VERSION AX172880.1 GI:14597915
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 4144)
AUTHORS Lu,P., Garman,J.D. and Candia,A.F.
TITLE Clasp-3 transmembrane protein
JOURNAL Patent: WO 0142297-A 7 14-JUN-2001;
Arbor Vita Corporation (US)
FEATURES
source Location/Qualifiers
1. 4144
/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="preliminary CLASP-3 cDNA"
<2. .3610
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC43727.1"
/db_xref="GI:14597916"
/translation="RVNRSRLSNNSNPDISGTPSPDDEVRSIIGSKGLDRSNSWNT
GGPKAPWGSNPSAESIQMDRSCNRMSSHTETSFLQTLTGRLP TKKLFHEELAL
QWVCSGSVRESALQQAWEFFELMVKSMVHLLYFNKLEAPKRSFRPERFMDIALY
STIASDIVSRFOKDTEMVERLNTSLAFELNDLSYMDRGFVFSLIKSCYKOVSSKLYS
LPNPSVLVSLRDLFLRITCSHEHYVTNLPCSLTPPASPSPVSSATSQSSGFSTNV
QDOKIANNEELSVPRQOHVLAGLVLELAVIILDPAGELGLEHKVIMVHNLSSH
DSDPRYSDPQIKARVAMLYLDLIGIIMETVPOLYDFTETHNORGRPICIAVDYES
GSMISQTVAMAIAGTSVPQLTRPGSFLLTSTSGRQHTPSAESSRLICLLMWLKNA
DETVLQKWFDTLSVLQNLRLDLVLCVSCFEYKGVKEFERMNSLTFFKSKDMRAKLE
EALIGSIGARQEMVRSRGQLERSPSGSAFGSQENLRMRKDMTHWRQNTKLDKRAE
IEHEALIDGNLATEANLILDLLEIVQTVSVTESKESLIGVLKVLHSMACNQSAV
YLQHCFA TORALVSKFPELLFEETEQCADLCRLRHCSSTIGTIRSHPSASLYLLM
RONFEIGNNFARVKMOVPMSSLSIVGTSQNFNEEPLRSIKTLITYAEEDLERETTF
PDQYQDVLVFNLMILSDITYAKKEHQEDPEMLIDLTRYIAKYQTSPODLRLTWQNMAG
KHSERNHAEEAQCIVHSALVAYEYLSMLEDRKYLPYGCVTFQNISSNVLESAYSDD
VVSPEEGICSGKYFTESGLVGLLEQAAAFSMAGYEAVNEVYKVLPIPIHEANRDAK
KLSTIHGKLQEAFSKIVHQSGWERMFGTYFRVGFYGTKEGDLDEQEFVYKEPAITKL
AEISHRLEGEYGERFGEDEVVEYIKDSNPVDKCKLPENKAYIQITYVEPYFDIYEMKDR
ITYFDKNYNLRFMYCPTPTLIDGRAHGELEHQEFKRTILTTSHAFPYIKTRVNTHKE
EILPTPIEVAIEDMOKKTOELAFATHODPADPKMLQMVLOQSVGTTVNOGLEVAQVF
LSETPSDPKLFRHNKRLRCLCFKDFTRCEDALRNKNSLIGPVQKEYQRELKLSPP"

BASE COUNT 1252 a 829 c 890 g 1173 t
ORIGIN

alignment_scores:
Quality: 6157.50 Length: 1201
Ratio: 5.131 Gaps: 1
Percent Similarity: 99.917 Percent Identity: 99.750

alignment_block:
US-09-737-246-2 x AX172880 ..

Align seg 1/1 to: AX172880 from: 1 to: 4144

891 LeuAsnArgSerArgSerLeuSerAsnSerAsnProAspIleSerGlyTh 907
:::|||||
5 GTAAATCGTTCTCGAAGCCTTAGTAATAGCAATCCAGATATATCTGGGAC 54
907 rProThrSerProAspAspGluValArgSerIleIleGlySerLysGlyL 924
|||||
55 TCCACAGTCACACAGATGATGAAGTTCGATCAATCATCGGAGTAAGGTT 104
924 euAspArgSerAsnSerTrpValAsnThrGlyGlyProLysAlaAlaPro 940
|||||
105 TAGATCGCTCCAATTCCTGGGTTAACACGTGCTGCCAAAAGCTGCCCA 154
941 TrpGlySerAsnProSerProSerAlaGluSerThrGlnAlaMetAspAr 957
|||||
155 TGGGATCCAACCCAGTCCAGAGTCAGAAATCAACACAGAGCTATGATCG 204
957 gSerCysAsnArgMetSerSerHisThrGluThrSerSerPheLeuGlnT 974
|||||
205 AAGTTGTAATCGTATGTCTTCGCACACAGAGACGTCAGAGTTTCTTACAAA 254
974 hrLeuThrGlyArgLeuProThrLysLysLeuPheHisGluGluLeuAla 990
|||||
255 CATTAACGGGACGCTTACCAACTAAAGCTTTTTCACGAGAGAGCTGGCT 304

991 LeuGI nTrpValValCysSerGlySerValArgGluSerAlaLeuGlnGI 1007
|||||
305 TTGCAGTGGGTGTTTGCAAGTGGCAGCGTTCGGGAATCAGCTTTGCACACA 354
1007 nAlaTrpPhePhePheGluLeuMetValLysSerMetValHisHisLeuT 1024
|||||
355 AGCCTGGTCTTTTGTGAATTAAATGTTAAAGACATGGTGACCATTTTAT 404
1024 y rPheAsnAspLysLeuGluAlaProArgLysSerArgPheProGluArg 1040
|||||
405 ACTTTAATGATAAACTTGAGGCTCCAAAGAAAAGTCGTTTCCAGAACGT 454
1041 pheMetaspPileAlaAlaLeuValSerThrIleAlaSerAspIleVal 1057
|||||
455 TTCATGGATGACATGGACCTCTGTGCAGCAGCATTTGCTAGTATATAGT 504
1057 IserArgPheGlnLysAspThrGluMetValGluArgLeuAsnThrSerL 1074
|||||
505 TTCACGATTTCAAGAGACACAGAATGGTTGAGAGACTCAATACAGACC 554
1074 euAlaPhePheLeuAsnAspLeuLeuSerValMetaspArgGlyPheVal 1090
|||||
555 TTGCATTTCTTCTCAATGATCTGTGTCTGTATGACACAGAGATTGT 604
1091 pheSerLeuIleLysSerCysTyrLysGlnValSerSerLysLeuTyrSe 1107
|||||
605 TTTAGCCTTATAAAGTCTGCTATAAACAAGGIGTCTTCAAAAGCTTTACTC 654
1107 rLeuProAsnProSerValLeuValSerLeuArgLeuAspPheLeuArgI 1124
|||||
655 ATTACCGAATCCAGTGTCTGTGTCCTTGAGGCTGGATTCTTCTACGAA 704
1124 leIleCysSerHisGluHisTyrValThrLeuAsnLeuProCysSerLeu 1140
|||||
705 TCATCTGCAGTCATGAGCACTATGTTACATTAAACTTACCCTGCAGCTTA 754
1141 LeuThrProProAlaSerProSerProSerValSerSerAlaThrSerGI 1157
|||||
755 CTTACTCCACCTGCATCTCCATCACCTTCTGTCTTCTGCACATCTCA 804
1157 nSerSerGlyPheSerThrAsnValGlnAspGlnLysIleAlaAsnMetP 1174
|||||
805 GAGTTCTGGATTCTTCTACGAATGTACAAGACCAGCAATGTCAAATATGT 854
1174 heGluLeuSerValProPheArgGlnGlnHisTyrLeuAlaGlyLeuVal 1190
|||||
855 TTGAATTATCCGTGCTTCCTCCGCCAACAGCATTTATTTGGCAGGACTTGTG 904
1191 LeuThrGluLeuAlaValIleLeuAspProAspAlaGluGlyLeuPheGI 1207
|||||
905 TTACAGAGCTGGCTGTCAATTTTAGACCCTGATGCTGAAGGACTGTTTG 954
1207 yLeuHisLysLysValIleAsnMetValHisAsnLeuLeuSerSerHisA 1224
|||||
955 ATTGCATAAGAAAGTCATCATATATGTTACACAAATTTACTCTCCAGTCACG 1004
1224 spSerAspProArgTyrSerAspProGlnIleLysAlaArgValAlaMet 1240
|||||
1005 ACTCAGACCCCGGCTACTCTGACCCCTCAGATTAAGGCTCGAGTGCCATG 1054
1241 LeuTyrLeuProLeuIleGlyIleIleMetGluThrValProGlnLeuTy 1257
|||||
1055 TTGTATCTACCTCTGATTGGTATTTATCATGGAACGTGTACTCAGCTGTA 1104
1257 rAspPheThrGluThrHisAsnGlnArgGlyArgProIleCysIleAlaT 1274
|||||
1105 TGATTTTACAGAAACTCACAAATCAACGAGGAAGACCAATTGTATAGCCA 1154
1274 hrAspAspTyrGluSerGluSerGlySerMetIleSerGlnThrValAla 1290
|||||
1155 CTGATGATTTATGAAAGTGAGAGCGGAAGTATGATAAGCCAGACCGTTGCC 1204

1291 MeAlaIleAlaGlyThrSerValProGlnLeuThrArgProGlySerPh 1307
|||||
1205 ATGGCAATCCGACGAGGACATCGGTCCCTCAACTAACAAGGCCCTGGCAGTTT 1254
1307 eLeuLeuThrSerThrSerGlyArgGlnHisThrThrPheSerAlaGluS 1324
|||||
1255 CCTCCTCACGTCACAGAGTGGCAGGCCAACACACTACCTTTTCAGCAGAAT 1304
1324 e rSerArgSerLeuLeuIleCysLeuLeuTrpValLeuLysAsnAlaAsp 1340
|||||
1305 CAAGTCGAAGCCTTTTGATCTGTCTACTTTGGGTTCTCAAAAATGCAGAT 1354
1341 GluThrValLeuGlnLysTyrPheThrAspLeuSerValleuGlnLeuAs 1357
|||||
1355 GAAACAGTTCTACAGAAAGTGGTTTACAGATCTCTCAGTCTTGCAAGCTAAA 1404
1357 nArgLeuLeuAspLeuLeuTyrLeuCysValSerCysPheGluTyrLysG 1374
|||||
1405 CCGGCTATAGATCTGCTTATCTCTGTGTCTTGTGCTTGAAGTATAAAG 1454
1374 IyLysLysValPheGluArgMetAsnSerLeuThrPheLysSerLys 1390
|||||
1455 GGAAAAAGTGTTTGAACGAATGATAGCTTGACCTTTAAAGAAATCAAAA 1504
1391 AspMetArgAlaLysLeuGlnGluAlaIleLeuGlySerIleGlyAlaAr 1407
|||||
1505 GACATGAGGCCAAAGCTTGAAGAGCTATTTCTGGAGCATAGGTGCCAG 1554
1407 gGlnGluMetValArgArgSerArgGlyGlnLeuGluArgSerProserG 1424
|||||
1555 GCAAGAAATGGTACGGCGAAGCCGAGACAGCTCGAGAGAAAGCCCATCTG 1604
1424 LySerAlaPheGlySerGlnGluAsnLeuArgTyrPArgLysAspMetThr 1440
|||||
1605 GAAGTGCCCTTTGGAAGTCAAGAAAAATTTGAGGTGGAGGAAGATATGACT 1654
1441 HisTrpArgGlnAsnThrGluLysLeuAspLysSerArgAlaGluIleGI 1457
|||||
1655 CACTGGCGTCAAAAACACAGAGAGAGCTTGACAANAATCAAGACAGAGATTGA 1704
1457 uHisGluAlaLeuIleAspGlyAsnLeuAlaThrGluAlaAsnLeuIle 1474
|||||
1705 ACACGAAGCACTGATTGATGGAACCTGGCTACAGAAAGCAAACTTAATCA 1754
1474 leLeuAspThrLeuGluIleValValGlnThrValSerValThrGluSer 1490
|||||
1755 TTTTAGATACATTAGAGATTGTGTTCAAGACCGTTTCTGTAACGGAATCC 1804
1491 LysGluSerIleLeuGlyGlyValLeuLysValLeuLeuHisSerMetAl 1507
|||||
1805 AAAGAGAGCATTTCTTGCTGGAGTGTCTTAAAGTGCTACTACACAGCATGGC 1854
1507 acCysAsnGlnSerAlaValTyrLeuGlnHisCysPheAlaThrGlnArgA 1524
|||||
1855 CTGTAACCAAAGTGCAGTTTATCTACAACAACCTGTTTGTGTACACAGAGAG 1904
1524 lAlaLeuValSerLysPheProGluLeuLeuPheGluGluGluThrGluGln 1540
|||||
1905 CCTTGGTTTCAAAAGTTTCTGAACTCTTATTGGAAGAAGAGACAGAGCAG 1954
1541 CysAlaAspLeuCysLeuArgLeuLeuArgHisCysSerSerSerIleGI 1557
|||||
1955 TGTGCTGATTATGCTCAGGCTTCTCCGACACTGTAGCAGTAGCATCCG 2004
1557 yThrIleArgSerHisProSerAlaSerLeuTyrLeuLeuMetArgGlnA 1574
|||||
2005 TACAAATACGGTCAACACCCAGTGCCTTCCCTTACTACTAATGAGGCCAAA 2054
1574 snPheGluIleGlyAsnAsnPheAlaArgValLysMetGlnValPromet 1590
|||||
2055 ACTTTGAGATTGGGAATAACTTTGCCAGGGGTTAAAAATGCAGGTACCAATG 2104
1591 SerLeuSerSerLeuValGlyThrSerGlnAsnPheAsnGluGluPheLe 1607

|||||
2105 TCACTATCCCTCGGTGGGCACATCTCAGAAATTTTAATGAGAAATTCCTT 2154
1607 uArgArgSerLeuLysThrIleLeuThrTyraIaGluAspLeuGluL 1624
2155 AAGACGTTCTTAAGACTATATGACATATGCTGAAGAGATCTGGAAT 2204
1624 euArgGluThrThrPheProaspGlnValGlnAspLeuValPheasnLeu 1640
2205 TGAGGGAACAACATTTCCGTGATCAGGTCAGGATCTGGTTTCAATCTC 2254
1641 HisMetIleLeuSerAspThrValLysMetLysGluHisGlnGluAspPr 1657
2255 CATATGATCTTTCTGATACTGTGAANAATGAAGGAACACCAGAGGATCC 2304
1657 oGluMetLeuIleaspLeuMetTyrArgIleAlaLysGlyTyrGlnThrS 1674
2305 TGAATGTGATGATCTAATGTACAGAAATGCCAAGGGTTACCAAGACCT 2354
1674 erProglu...ArgLeuThrTrpLeuGlnasnMetAlaGlyLysHisSer 1689
2355 CTCAGATCTGCGATGTGACCTGTTGCAAGACATGGCAGCAAGCACTCA 2404
1690 GluArgSerAsnHisAlaGluAlaAlaGlnCysLeuValHisSerAlaAl 1706
2405 GAACGAAGCAATCATGCTGAAGCTGCACAGTGTCTAGTCCACCTCAGAC 2454
1706 aleuValAlaGluTyrLeuSerMetLeuGluAspArgLysTyrLeuPro 1723
2455 ACTTGTGCTGAATATTGAGCATGCTGGAGGACCGGAATATCTTCCCTG 2504
1723 aIGLycysValThrPheGlnasnIleSerSerasnValLeuGlnGluSer 1739
2505 TGGGATGTGTAACATTTCAGAATATTCATCTAATGTTTAGAAGAAATCT 2554
1740 AlaValSerAspAspValValSerProaspGlnGluGlyIleCysSerG 1756
2555 GCGGTCTCAGATGATGTGTATCTCCAGATGAAGAGGTATCTGCTCTGG 2604
1756 yLysTyrPheThrGlnSerGlyLeuValGlyLeuGlnAlaAla 1773
2605 AAAATACTTTACTGAGTCAGGACTGTGGCATTTACTGGAACAAGCAGCTG 2654
1773 laseRPheserMetAlaGlyMetTyrGluAlaValAsnGluValTyrLys 1789
2655 CTTCTTCTGTATGGCTGGCATGTATGAAGCAGTAATGAAGTTTACAA 2704
1790 ValLeuIleProIleHisGluAlaAsnArgAspAlaLysLysLeuSerTh 1806
2705 GTACTTATTCCTATTCATGAAGCTAATCGGGAATGCAAGAAGAACTATCCAC 2754
1806 rIleHisGlyLysLeuGlnGluAlaPheSerLysIleValHisGlnSert 1823
2755 AATTCATGGTAACCTTCAGAAGCAATTCAGCAAAATGTTTCATCAGAGTA 2804
1823 hrGlyTyrGluArgMetPheGlyThrTyrPheArgValGlyPheTyrGly 1839
2805 CTGGCTGGAGCGGATGTTGGCACCTATTTTGGTGTGTTTATGGA 2854
1840 ThrLysPheGlyAspLeuAspGlnGlnIupheValTyrLysGluProAl 1856
2855 ACCAAGTTGCGGATTTGGATGAACAAGAAATTTGTTACAAAGGAGCCCTGC 2904
1856 aIleThrLysLeuAlaGluIleSerHisArgLeuGlnGlyPheTyrGlyG 1873
2905 AATAACCAAACTTCAGAGATATCTCACAGATTGGAGGGATTTTACGGAG 2954
1873 luArgPheGlyGluAspValValGluValIleLysAspSerasnProVal 1889
2955 AAAGATTTGGAGAGGATGTGTTGAAGTAATCAAAAGACTCTAATCCTCTGA 3004
1890 AspLysCysLysLeuAspProAsnLysAlaTyrIleGlnIleThrTyrVa 1906
|||||

3005 GACAAGTGTAATTAGATCTTAACAAGGCATATATTCAGATTACCTAGT 3054
1906 lGluProTyrPheAspThrTyrGluMetLysAspArgIleThrTyPhea 1923
3055 GGAGCCATACTTTGACACATATGAGATGAAGGAGACAATCACCTATTTGC 3104
1923 sPlysAsnTyrAsnLeuArgArgPheMetTyrCysThrProPheThrLeu 1939
3105 ACAAAAATTACAATCTTCGTGCGATTGATGATGATGATGATGATGATGAT 3154
1940 AspGlyArgAlaHisGlyGluLeuHisGlnGlnPheLysArgLysThrI 1956
3155 GATGGCCGTGCCCATGGGGAACCTTCATGAACAATTCAAAAGAGACCAT 3204
1956 eleuThrThrSerHisAlaPheProTyrIleLysThrArgValAsnValT 1973
3205 TCTGACTACGTCATGCTGCTTCTTCTTATATTAACAAGGGTCAATGTCA 3254
1973 hrHisLysGluGluIleIleLeuThrProIleGluValAlaIleGluAsp 1989
3255 CTCATTAAGAAGAGATCATCTTAACACCAATTGAAGTTGCTATTGAGGAC 3304
1990 MetGlnLysLysThrGlnGluLeuAlaPheAlaThrHisGlnAspProAl 2006
3305 ATGCAGAAAAAGACACAGAGATTGGCATTTGCCAACACATCAGATCCGCC 3354
2006 aaspProLysMetLeuGlnMetValLeuGlnGlySerValGlyThrThrV 2023
3355 AGACCCCAAAATGCTTCAGATGTGACTCCAGGATCTGTAGGCACACAG 3404
2023 alaSnGlnGlyProLeuGluValAlaIaGlnValPheLeuSerGluIlePro 2039
3405 TGAATCAGGGGCTTGGAAAGTTGCCAGGTTTCTGTCTGAATTAATCT 3454
2040 SeraspProLysLeuPheArgHisHisAsnLysLeuArgLeuCysPheLy 2056
3455 AGTGACCCAAAGCTTTCAGACATCATTAATACTGCGACTCTGCTTTAA 3504
2056 sasPheThrLysArgCysGluAspAlaLeuArgLysAsnLysSerLeuI 2073
3505 AGATTTTACTAAAGGTGTGAAGATGCCCTTAAGAAAAAATAAGAGCTTAA 3554
2073 leglyProValGlnLysGluTyrGlnArgGluLeuGlyLysLysSerSer 2089
3555 TTGGCCGGTTCAAAAGAGATATCAAAAGGGAATTGGGGAAGTAATCTTCG 3604
2090 pro 2090
3605 CCT 3607
seq_name: gb_pat:AX172932
seq_documentation_block: 4144 bp DNA linear PAT 03-JUL-2001
LOCUS AX172932
DEFINITION Sequence 59 from Patent WO0142297.
ACCESSION AX172932
VERSION AX172932.1 GI:14597917
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 4144)
AUTHORS Lu, P., Garman, J.D. and Candia, A.F.
TITLE Clasp-3 transmembrane protein
JOURNAL Patent: WO 0142297-A 59 14-JUN-2001;
Arbor Vita Corporation (US)
FEATURES
source 1..4144
/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="human CLASP-3 cDNA"
CDS 197..3610


```
/note="human CLASP-3"
/codon_start=1
/protein_id="CAC43728.1"
/db_xref="GI:14597918"
/translation="MDRSCNRMSSTHETSSFLQTLTGLPTKKLFHEELALOMVVCSC
SVRESALQOAMFEFFELMVKSMVHLYFNDKLEAPRKSRRPERFMDIALVSTIASDI
VSRFQKDTFEMVERLNTSLAFFLNDLSYMDRGFVSLKSCYKQVSSKLYSLPNPSYL
VSLRLDFLRILCSHEHYVTNLPCSLILTPPASPSPSVSSATSOSSGFSTNVQDKIAN
MEELSVPRQOHYLAGLVLTSLAVLIDDAEGLFKKVINMYHNLLSSHSDSPRYS
DPOIKARVAMLYPLIGIMETVPQLYDFTETHNQRGPICATDDYSESGSMISOT
VAMAIAGTSVPQLTRPGSFLTSTGRQHTFSAESSRLICLLWLKNADETVLOK
WFTDLSTVLQNRLLDLLYLCVSCFEYKQKVFERNMSTFKSKMDRAKLEAIIIGSI
GARQEMVRRSRGQLERSPSGSAFGSQENLRWKDMTHWRQNEKLDKSRAIEHEALI
DGNLATLEANLIIIDTLEIVVQTVSTESKESILGGVLKVLHSMACNQSAYVLIQCPA
TORALVSKPELLFEEETEQCADLCRLLRHCSSSIGTIRSPSASYLLMRQNEIG
NNFARVKMOVMSLSSLVGTSQNFNEFLRSLKTLITVAEEDLELRETFPDQVODL
VENLHMLSDTVKMKHEQEDPEMLIDLMYRIAKGYQTSPLRLTLWLQNMAGKHSERN
HAEAAOCLVHSAALVAEYLSMLEDRKYLPGCVTFQNISSNVLESASVSDVVSPEE
GICSGKYFTESGLVGLLEQAAASESMAGMYEAVNEVYKLIPIHEANRDAKKLSTIHG
KLQEAFSKIYHQSTGWMERMGFTYFRVGYGTFGDLDEQEEFYKKEPAITKLAISHRL
EGFYGERFGEDEVVEVIKDSNPVDCKLDPNKAYIOITYVEPYFDTYEMKDRITYFDKN
YNLRFEMYCPTFTLDGRAHGEIHEQEKRTILTTSHPYIKTRVNVTHKEIILTPI
EVAIEDMOKRQELAFATHODPADPKMLQMVLOGSVGTTVNOGLEVAQVFLSEIPSD
PKLFRHNKLRLCFKDFTRKCEDALRKNSLIGPVQKEYQRELKLISSP"

BASE COUNT      1252 a      829 c      890 g     1173 t
ORIGIN

alignment_scores:
      Quality: 6157.50,      Length: 1201
      Ratio: 5.131      Gaps: 1
Percent Similarity: 99.917      Percent Identity: 99.750

alignment_block:
US-09-737-246-2 x AXI72932 ..

Align seg 1/1 to: AXI72932 from: 1 to: 4144

891 LeuAsnArgSerArgSerLeuSerAsnSerAsnProAspIleSerGlyTh 907
:::|||||
5 GTAATCGTTCTCGAAGCCTTAGTAATAGCAATCCAGATATATCTGGGAC 54

907 rProThrSerProAspAspGluValArgSerIleIleGlySerIleGlyL 924
|||||
55 TCCACAGTCACACAGATGATGAAGTTCGATCAATCATCGGAGTAAGGTT 104

924 euAspArgSerAsnSerTrpValAsnThrGlyGlyProIleAlaAlaPro 940
|||||
105 TAGATCGGTCCCAATTCCTGGGTTAACACTGGTGTCCAAAGCTGCCCCA 154

941 TrpGlySerAsnProSerProSerAlaGluSerThrGlnAlaMetAspAr 957
|||||
155 TGGGATCCAAACCCCAAGTCCCAAGTGCAGAAATCAACACAGGCTATGATCG 204

957 gSerCysAsnArgMetSerSerHisThrGluThrSerSerPheLeuGlnT 974
|||||
205 AAGTTGTAATCGTATGTTCTTCGACACAGAGACGTCAAGTTTCTTACAAA 254

974 hrLeuThrGlyArgLeuProThrIleIleGlySerPheHisGluGluLeuAla 990
|||||
255 CATTAAAGGAGCGCTTACCAACTAAAAGCTTTTTCACAGAGAGCTGGCT 304

991 LeuGlnTrpValValCysSerGlySerValArgGluSerAlaLeuGlnG 1007
|||||
305 TTGCAGTGGGTGTTTTCAGTGGCAGCGCTTCGGGAATCAGCTTTCACAACA 354

1007 naLaTrpPhePhePheGluLeuMetValIleSerMetValHisHisLeuT 1024
|||||
355 AGCCTGGTCTTTTTCGAATTAATGTTAAAGACATGCTGCACCATTTAT 404

1024 yrPheAsnAspIleGluAlaProArgIleGlySerArgPheProGluArg 1040
|||||
405 ACTTTAATGATAAACTTGAGGCTCCAAAGAAAGTCTGTTTCCAGAACGT 454
```

```
1041 PheMetAspAspIleAlaAlaLeuValSerThrIleAlaSerAspIleVal 1057
|||||
455 TTCATGGATGACATTTGCAGCTCTTGTCAGACAGATTTGCTAGTATATAGT 504

1057 lSerArgPheGlnIleLysAspThrGluMetValGluArgLeuAsnThrSerL 1074
|||||
505 TTCACGATTTTCAGAAGGACACAGAAATGGTTGAGAGACTCATATACAAGCC 554

1074 euAlaPhePheLeuAsnAspLeuLeuSerValMetAspArgGlyPheVal 1090
|||||
555 TTGCATTCTTCTCAATGATCTGTGTCTGTATGACACAGAGATTGTT 604

1091 PheSerLeuIleLysSerCysTyrlYsglnValSerSerIleuTyrsE 1107
|||||
605 TTTAGCCTTATAAAGTCTGCTATTAACAGGTGCTTCAAGCTTTACTC 654

1107 rLeuProAsnProSerValLeuValSerLeuArgLeuAspPheLeuArgI 1124
|||||
655 ATTACCGAATCCCAAGTGTCTGCTGCTCTGAGGCTGGATTTCACGAA 704

1124 lIleIleCysSerHisGluHisTyrlValThrLeuAsnLeuProCysSerLeu 1140
|||||
705 TCATCTGCAGTCATGAGCACATATGTACATTAAACTTACCCTGCAGCTTA 754

1141 LeuThrProProAlaSerProSerProSerValSerSerAlaThrSerG 1157
|||||
755 CTTACTCCACCCTGCATCTCCATCACCTCTGTTCTTCTGACACATCTCA 804

1157 nSerSerGlyPheSerThrAsnValGlnAspGlnLysIleAlaAsnMetP 1174
|||||
805 GAGTTCGTGATTTCTACGAATGTACAGAACCAAAAGATTGCATAATATGT 854

1174 heGluLeuSerValProPheArgGlnGlnHisTyrlLeuAlaGlyLeuVal 1190
|||||
855 TTGAATTATCCGTCCTTTCGCCCAACAGCATTAATTGGCAGGACTTGTG 904

1191 LeuThrGluLeuAlaValIleLeuAspProAspAlaGluGlyLeuPheG 1207
|||||
905 TTAACAGAGCTGGCTGTATTTAGACCTGATGCTGAAGAGACTGTTGG 954

1207 yLeuHisIleLysValIleAsnMetValHisAsnLeuLeuSerSerHisA 1224
|||||
955 ATTGCATAGAAGATCATCATATATGGTACACAAATTACTCTCCAGTCACG 1004

1224 sPseAspProArgTyrlSerAspProGlnIleIleLysAlaArgValAlaMet 1240
|||||
1005 ACTCAGACCCCGGCTACTCTGACCCCTCAGATTAAGGCTCGAGTGGCCATG 1054

1241 LeuTyrlLeuProLeuIleGlyIleIleMetGluThrValProGlnLeuTy 1257
|||||
1055 TTGTATCTACCTCTGATTTGGTATTATCATGGAATACTGATCCTCAGCTGTA 1104

1257 rasPheThrGluThrHisAsnGlnArgGlyArgProIleCysIleAlaT 1274
|||||
1105 TGATTTTACAGAAACTCACAAATCAACGAGAGAACCAATTGTATAGCCA 1154

1274 hrAspAspTyrlGluSerGluSerGlySerMetIleSerGlnThrValAla 1290
|||||
1155 CTGATGATTATGAAGATGAGAGCGGAGATGATGATAAGCCAGACCGTTGCC 1204

1291 MetAlaIleAlaGlyThrSerValProGlnLeuThrArgProGlySerPh 1307
|||||
1205 ATGGCAATCGCAGGGACATCGGTCCCTCAACTAACAAAGCCGTGGCAGTTT 1254

1307 euLeuLeuThrSerThrSerGlyArgGlnHisThrThrPheSerAlaGlnS 1324
|||||
1255 CCTCCTCAGCTCAACGAGTGGCAGGCAACACACTACTCTTTTCAGCAGAAAT 1304

1324 erSerArgSerLeuLeuIleCysLeuLeuTrpValLeuLysAsnAlaAsp 1340
|||||
1305 CAAGTCGAAGCCTTTTGTATCTGTACTTTGGGTTCTCAAAAATGACAGAT 1354
```

1341 GluThrValLeuGlnLysTrpPheThrAspLeuSerValLeuGlnLeuAs 1357
|||||
1355 GAACAGCTTCTACAGAGTGGTTTACAGATCTCTCAGCTCTGCAGCTAA 1404
1357 naRgLeuLeuAspLeuLeuTyrLeuCysValSerCysPheGluTyrLysG 1374
|||||
1405 CCGGCTATTGATCTGCTTATCTCTGTGTGCTTGGCTTTGAGTATAAG 1454
1374 LylsLysValPheGluArgMetAsnSerLeuThrPheLysLysSerLys 1390
|||||
1455 GCAAAAAAGTGTGGAACGAATGAATAGCTTGACCTTTAAGAATCAAA 1504
1391 AspMetArgAlaLysLeuGluGluAlaIleLeuGlySerIleGlyAlaAr 1407
|||||
1505 GACATGAGAGCAACGCTTGAAGAAGCTATTCTTGGAGCATAGGTGCCAG 1554
1407 gGlnGluMetValArgArgSerArgGlyGlnLeuGluArgSerProSerG 1424
|||||
1555 GCAAGAAATGGTACGGCGAAGCCGAGAGACGCTCGAGAGAAGCCCATCTG 1604
1424 LysSerAlaPheGlySerGlnGluAsnLeuArgTrpArgLysAspMetThr 1440
|||||
1605 GAAGTGCCCTTGGAGTCAAGAAATTTGAGTGGAGGAAGATATGACT 1654
1441 HisTrpArgGlnAsnThrGluLysLeuAspLysSerArgAlaGluIleG 1457
|||||
1655 CACTGGCGTCAAAACACAGAGAAGCTTGACAAATCAAGAGCAGAGATTGA 1704
1457 uHisGluAlaLeuIleAspGlyAsnLeuAlaThrGluAlaAsnLeuIle 1474
|||||
1705 ACACGAAGCACGTGATTGATGGAACCTGGCTACAGAAGCAACCTATCA 1754
1474 IeLeuAspThrLeuGluIleValValGlnThrValSerValThrGluSer 1490
|||||
1755 TTTTAGATACATTAGAGATTGTGTTCAGACCGTTTCTGTAACCGCAATCC 1804
1491 LysGluSerIleLeuGlyGlyValLeuLysValLeuLeuHisSerMetAl 1507
|||||
1805 AAAGAGAGCATTTCTGGTGAGTGCTAAAAAGTGCTACTACACAGCATGGC 1854
1507 acYsAsnGlnSerAlaValTyrLeuGlnHisCysPheAlaThrGlnArgA 1524
|||||
1855 CTGTAACCAAGTGCAGTTTATCTACACACTGTTTGTACACAGAGAG 1904
1524 IaLeuValSerLysPheProGluLeuLeuPheGluGluIuThrGluGln 1540
|||||
1905 CCTTGGTTCAAGTTTCCCTGAACCTTATTTGAAGAAGAGACAGACAG 1954
1541 CysAlaAspLeuCysLeuArgLeuLeuArgHisCysSerSerSerIleG 1557
|||||
1955 TGTGCTGATTATGCTCAGGCTTCTCCGACACTGTAGCAGTAGCATCGG 2004
1557 YThrIleArgSerHisProSerAlaSerLeuTyrLeuLeuMetArgGlnA 1574
|||||
2005 TACAATACGGTCACACCCAGTGCCTCCCTTTACTACTAATGAGCAAA 2054
1574 snPheGluIleGlyAsnAsnPheAlaArgValLysMetGlnValProMet 1590
|||||
2055 ACTTGAGATTGGGAATACTTTGCCAGGGTTAAATGCAAGTACCCAATG 2104
1591 SerLeuSerSerLeuValGlyThrSerGlnAsnPheAsnGluPheLe 1607
|||||
2105 TCACTATCCTCCTTGGTGGGACACATCTCAGAAATTTAATGAAGATTCTT 2154
1607 uArgArgSerLeuLysThrIleLeuThrTyrAlaGluGluAspLeuGluL 1624
|||||
2155 AAGACGTTCTCTAAAGACTATATTGACATATGCTGAAGAAGATCTGGAAT 2204
1624 euArgGluThrThrPheProAspGlnValGlnAspLeuValPheAsnLeu 1640
|||||
2205 TGAGGGAACAACATTTCTGATCAGGTCACGATCTGGTTTCAATCTC 2254
1641 HisMetIleLeuSerAspThrValLysMetLysGluHisGlnGluAspPr 1657

|||||
2255 CATATGATTCTTCTGATACTGTGAAAATGAAGGAACACAGAGATCC 2304
1657 oGluMetLeuIleAspLeuMetTyrArgIleAlaLysGlyTyrGlnThrS 1674
|||||
2305 TGAATGTTGATTGATCTAATGTACAGAATTGCCAAGGTTACCAGACCT 2354
1674 eRProGlu...ArgLeuThrTrpLeuGlnAsnMetAlaGlyLysHisSer 1689
|||||
2355 CTCAGATCTCGGATTCGACCTGGTTGCAGAACATGGCAGGCAAGCACTCA 2404
1690 GluArgSerAsnHisAlaGluAlaAlaGlnCysLeuValHisSerAlaAl 1706
|||||
2405 GAACGAAGCAATCATGCTGAGCTGCACAGTGTCTAGTCCACTCAGCAGC 2454
1706 aLeuValAlaGluTyrLeuSerMetLeuGluAspArgLysTyrLeuPro 1723
|||||
2455 ACTTGTTCCTGAATATTGTGACCATGCTGGAGGACCGGAATATCTTCCTG 2504
1723 aGlyCysValThrPheGlnAsnIleSerSerAsnValLeuGluGluSer 1739
|||||
2505 TGGGATGTGTAAACATTTCAGAAATATTTCAATCTAATGTTTGAAGAATCT 2554
1740 AlaValSerAspAspValValSerProAspGluGluGlyIleCysSerG 1756
|||||
2555 GCGGTCTCAGATGATGTGTATCTCCAGATGAAGAAGTATCTGCTCTGG 2604
1756 YLysTyrPheThrGluSerClyLeuValGlyLeuLeuGluGlnAlaAla 1773
|||||
2605 AAATACTTTTACTGAGTCAGGACTTGTGGGATTACTGGAACAAGCAGCTG 2654
1773 IAserPheSerMetAlaGlyMetTyrGluAlaValAsnGluValTyrLys 1789
|||||
2655 CTTCCTTCTCTATGGCTGCGCATGTATGAAGCAGTTAATGAAGTTTACAA 2704
1790 ValLeuIleProIleHisGluAlaAsnArgAspAlaLysLysLeuSerTh 1806
|||||
2705 GTRACTATTCTTATTCATGAAGCTAATCGGGATGCCAAAGAACTATCCAC 2754
1806 rIleHisGlyLysLeuGlnGluAlaPheSerLysIleValHisGlnSert 1823
|||||
2755 AATTCATGTTAAACTTCAAGAAGCATTCAGCAAAATTTTCATCAGAGTA 2804
1823 hRgLYTrpGluArgMetPheGlyThrTyrPheArgValGlyPheTyrGly 1839
|||||
2805 CTGGCTGGGAGCGGATGTTGGCACCTATTTCGTGTGTTTATGCA 2854
1840 ThrLysPheGlyAspLeuAspGluGlnIuPheValTyrLysGluProAl 1856
|||||
2855 ACCAAGTTCGGGATTTGGATGAACAAGAATTTGTTACAAGAGCCCTGC 2904
1856 aIleThrLysLeuAlaGluIleSerHisArgLeuGluGlyPheTyrGlyG 1873
|||||
2905 AATAACCAAACTTGCAGAGATATCTCACAGATTGGAGGATTTTACGGAG 2954
1873 LuArgPheGlyGluAspValValGluValIleLysAspSerAsnProVal 1889
|||||
2955 AAAGATTTGGAGAGGATGTGTTGAAGTAATCAAGAAGCTTAATCCTGTA 3004
1890 AspLysCysLysLeuAspProAsnLysAlaTyrIleGlnIleThrTyrVa 1906
|||||
3005 GACAAAGTGTAAATTAGATCCTAACCAAGGCATATATTAGATTACCTATGT 3054
1906 IgluProTyrPheAspThrTyrGluMetLysAspArgIleThrTyrPhea 1923
|||||
3055 GGAGCCATACTTTGACACATATGAGATGAAGGACAGAATCACCTATTTCG 3104
1923 sPLysAsnTyrAsnLeuArgArgPheMetTyrCysThrProPheThrLeu 1939
|||||
3105 ACAAAAATTACAAATCTTCGTCGATTCATGTACTGTACACCCCTTACTTTA 3154
1940 AspGlyArgAlaHisGlyGluLeuHisGluGlnPheLysArgLysThrIl 1956
|||||


```

162 luAlaProAspGlyAsnSerTyrGlnAspAspGlnAspAspLeuTyrArg 178
407 CCTTGAG...TGCAGTGAACCCGCTGCTCAGGCAGGCCCGCCAC 450
179 ArgSerMetSerIleAspAspThrProArgGlySerTirPalAcysSerI 195
451 TTAACGTGCTGTGCGACGCTGTCTGGGAAGGCCCGCTCAGCTGCTGTGA 500
195 ePheAspLeuLysAsnSerLeuProAspAlaLeuLeuProAsnLeuLeu 212
501 CTTTGACCTCCGCAGCCTGCAGCCTGTACAAAGCGGCTAGAAAACCTCTGC 550
212 spArgThrProAsnGluGluIleAspArgGlnAsnAspAspGlnArgLys 228
551 AGCAAGTGAAGTCCGAGAGACTTTGAAAGCAGAACGAGAGAGGCCCGGAG 600
229 SerAsnArgHisLysGluLeuPheAlaLeuHisProSerProAspGluG 245
601 ACCAACAAGCAGGCGCCGAGCTCTTGCCCTTACCCATCAGTGGACAGAGA 650
245 uGluProIleGluArgLeuSerValProAspIleProLysGluHisPheG 262
651 GGATGCTGTGGAATACGTCCAGTACCAAGATGTCACCAAGGAACACTGG 700
262 LysGlnArgLeuValLysCysLeuSerLeuLysPheGluIleGluIle 278
701 GCACAGAAATATTGGTCAAGTTGCTGACCTTGAAGTTGAGATTGAAATT 750
279 GluProIlePheAlaSerLeuAlaLeuTyrAspValLysGluLysLys 295
751 GAGCCCTGTTGCCAGCATTGCCCTTACGATGTTAAGAAAGAAAGAAA 800
295 sIleSerGluAsnPheTyrPheAspLeuAsnSerGluGlnMetLysGly 312
801 GATCTCAGAAAATTTTCACTGTGACCTGAACTCTGACCAGTTCAAGGAT 850
312 euLeuArgProHisValProProAlaAlaIleThrThrLeuAlaArgSer 328
851 TTCTGCGAGCTCACACGCCCTTCACTGCGCCGATCAAGTCAAGCGAGATCT 900
329 AlaIlePheSerIleThrTyrProSerGlnAspValPheLeuValIle 345
901 GCAGTCTTCTCAGTACACTACCCCTCTCAGACATCTACTGTTAGTCAA 950
345 sLeuGluLysValLeuGlnGlnLysAspIleGlyGluCysAlaGluPro 362
951 GATTGAAAAAGTCCCTGCAGCAGGAGATATTGGAGACTGTGCAGACCCCT 1000
362 yMetIlePheLysGluAlaAspAlaThrLysAsnLysGluLysLeuGlu 378
1001 ACACGGTTATCAAGAAAGTGTGTGGAAGAGTAAAGAAAGATTGAA 1050
379 LysLeuLysSerGlnAlaAspGlnPheCysGlnArgLeuGlyLysTyrAr 395
1051 AAACATAAACTCCAAGCTGAATCTTCTGACCAGCGTTTGGGGAATAACCG 1100
395 gMetProPheAlaThrThrAlaIleHisLeuMetAsnIleValSerSera 412
1101 GATGCCCTTTGGCTGGGCACCACTAAAGCTTATCAAGCTTCTTC...AATG 1147
412 laGlySerLeuGluArgAspSerThrGluValGluIleSerThrGlyGlu 428
1148 TCTCCACCCCTTGAGAGGAGGAGTAAGTATGATGTGACTCTGTGTTGGG... 1194
429 ArgLysGlySerTyrSerGluArgArgAsnSerSerIleValGlyArgAr 445
1195 ...AGAAAGCCAGTGGGTGAACGAGAGACATTGGCC...CAATCTAGAAG 1238
445 gSerLeuGluArgThrThrSerGlyAspAspAlaCysAsnLeuThrSerP 462
1239 GCTTCTGAAAGAGCCCTCTCTCTTGAGAGGAAATGGGGTTGGATCCAAC 1288
462 heArgProAlaThrLeuThrValThrAsnPhePheLysGlnGluGlyAsp 478

```

[illegible]

2189 ATCAGAAAATCAGCGAGATGGCGCTGGAGCATGAGCTGAAGCTCAGCATC 2238
775 SerAlaLeuAsnSerSerGlnLeuGluProValValArgPheLeuHisLe 791
:::|||||:::|||||:::|||||:::|||||:::|||||
2239 ATCTGCCTGAACTCCTCCCGCCTGGAGCCGCTGCTCTCTCCCTGCACCT 2288
791 uLeuLeuAspLysLeuIleLeuLeuValIleArgProProValIleAlaG 808
|:::|||||:::|||||:::|||||:::|||||:::|||||
2289 GGTGCTGGACAAGCTCTTCCAGCTGTCCGTGCAGCCCATGGTCATCGCTG 2338
808 lYglnIleValAsnLeuGlyGlnAlaSerPheGluAlaMetAlaSerIle 824
|||||:::|||||:::|||||:::|||||:::|||||
2339 GCCAGACAGCCAACTTCTCCAGTTTGCCTGCAGTCCGTGGTGGCCATC 2388
825 IleAsnArgLeuHisLysAsnLeuGluGlyAsnHisAspGlnHisGlyAr 841
||| |:::|||||:::|||||:::|||||:::|||||
2389 GCCAAACAGTCTGCACACACAGCAAGACCTGAGCAAGACCATGCGAG 2438
841 gasnSerLeuAlaSerTyrIleHisTyrValPheArgLeuProAsnT 858
||||| |:::|||||:::|||||:::|||||:::|||||
2439 GAAC TGCCCTGCTGCTGCTCCTACGTGACACTACGTCTCCGCGCTGCCAGAGG 2488
858 hTyrProAsnSerSerSerProGly..ProGlyGlyLeuGlyGlySer 873
:::|||||:::|||||:::|||||:::|||||
2489 TGCAAAAGGATGTGCCCAAGTCAAGCGCTCCACTGCCCTCTCAGACCCCT 2538
874 ValHisTyrAlaThrMetAlaArgSerAlaValArgProAlaSerLeuAs 890
||| |:::|||||:::|||||:::|||||:::|||||
2539 CGGAGCTTACCACACGTATGGCCGCACATCAGCTGCTGTGAGTTCAAA 2588
890 nLeuAsnArgSerArgSerLeuSerAsnSerAsnProAspIleSerGlyT 907
:::|||||:::|||||:::|||||:::|||||:::|||||
2589 GCTGCTGCAGGCCCGGTGATGAGCAGCAGTAACCCAGACCTCGCGGGGA 2638
907 hTyrThrSerProAspAspGluValArgSerIleIleGlySerLysGly 923
||| |:::|||||:::|||||:::|||||:::|||||
2639 CACACTCCGCAGCAGACGAGGAAGTGAAGAACAATCATGCTTCAAAAG... 2685
924 LeuAspArgSerAsnSerTrpValAsnThrGlyGlyProLysAlaAlaIpr 940
2685 2685
940 cTrrpGlySerAsnProSerProSerAlaGluSerThrGlnAlaMetAspa 957
||| |
2686ATCGCCGATC 2695
957 rgSerCysAsnArgMetSerSerHisThrGluThrSerSerPheLeuGln 973
||:|||||:::|||||:::|||||:::|||||
2696 GCAACTGCAGCCGAATGCTTACTATGTGCTGCGCAGTAGTAGTGCACA 2745
974 Thr.....LeuThrGlyArgLeuProThrLysLysLeuPheHisGluG 988
:::|||||:::|||||:::|||||:::|||||
2746 AGTTCACCTGCAGCCCAAGCGCCAGCCAGCAAAAGCATTTCCATGAGGA 2795
988 uLeuAlaLeuGlnTrpValValLysSerGlySerValArgGluSerAlaL 1005
||||| |:::|||||:::|||||:::|||||
2796 GCTTGCCCTTCAGATGGTGGTCAACACCGGA..... 2826
1005 euGlnGlnAlaTrpPhePheGluLeuMetValLysSerMetValHis 1021
||| |:::|||||:::|||||:::|||||
2827ATGGTGAAGCAATGGCCACG 2847
1022 HisLeuTyrPheAsnAspLysLeuGluAlaProArgLysSerArgPhePr 1038
||| |:::|||||:::|||||:::|||||
2848 CACGTACATTAACATGACAAACGGGACAGTTTTCGAGGACTCGTTTYTTC 2897
1038 oGluArgPheMetAspIleAlaAlaLeuValSerThrIleAlaSerA 1055
:::|||||:::|||||:::|||||:::|||||
2898 TGACCGTTTCATGATGACATTAAC TACTATGTTTAATGTGGTCACCTCGG 2947
1055 spIleValSer.....ArgPheGlnLysAspThrGluMetValGlu 1068
:::|||||:::|||||:::|||||:::|||||
2948 AAATTGCAGCCCTTTTAAATAAACACACAGAAAGAAATGAACAGGCGGAA 2997

1069 ArgLeuAsnThrSerLeuAlaPhePheLeuAsnAspLeuLeuSerValMe 1085
:::||||| |:::|||||:::|||||:::|||||
2998 AAGATGAACATCAGCCTGGCTTCTCTGTATGACCTCTCTCCCTCAT 3047
1085 tAspArgGlyPheValPheSerLeuIleLysSerCysTyrLysGlnValS 1102
|:::|||||:::|||||:::|||||:::|||||
3048 GGATCGGGGCTTGTGTTAACTCATCAGACATTTATTCAGCCAGCTGT 3097
1102 eSerLysLeuTyrSerLeuProAsnProSerValLeuValSerLeuArg 1118
||:|||||:::|||||:::|||||:::|||||
3098 CAGCCAAAGCTCAGTAACCTTCCA.....ACGCTCATTTCCATGAGG 3138
1119 LeuAspPheLeuArgIleIleCysSerHisGluHisTyrValThrLeuAs 1135
||:|||||:::|||||:::|||||:::|||||
3139 CTAGAGTTCTTGAGAACTCTGTAGCCATGAGCATTAACCTCAATCTGAA 3188
1135 nLeu...ProCysSerLeuLeuThrProAlaSerProSerProServ 1151
||| |:::|||||:::|||||:::|||||
3189 CCTTTTATTATGAATGCTGATGCTGCCAACATCTCTTGCTTCCA 3238
1151 aLerSerAlaThrSerGlnSerSerGlyPheSerThrAsnValGlnAsp 1167
:::|||||:::|||||:::|||||:::|||||
3239 TATCT.....TCCCAAGACTCAAGCTCCTGCTCCAGCTTCCAGGAC 3279
1168 GlnLysIleAlaAsnMetPheGluLeuSerValProPheArgGlnGlnHi 1184
|:::|||||:::|||||:::|||||:::|||||
3280 CAGAGATCGCCACGATGTTGCATCTGACTTCCGAGTACCGCCACGACGA 3329
1184 sTyrLeuAlaGlyLeuValLeuThrGluLeuAlaValIleLeuAspProA 1201
|:::|||||:::|||||:::|||||:::|||||
3330 CTTCCTTCACCGGGCTCCTCTTTCACAGAACTGGCTGCTGCCCTGGATGCCG 3379
1201 spAlaGluGlyLeuPheGlyLeuHisLysLysValIleAsnMetValHis 1217
:::|||||:::|||||:::|||||:::|||||
3380 AAGGGAGGAATCAGCAAAAGTACAAAGAAAGAGTGCAGTAATTCCAC 3429
1218 AsnLeuLeuSerSerHisAspSerAspProArgTyrSerAspProGlnI 1234
:::|||||:::|||||:::|||||
3430 AGCCTGCTAAGTTCTCAGACCTGGACCCACGCTGTGTCAAAACGAGGT 3479
1234 eLysAlaArgValAlaMetLeuTyrLeuProLeuIleGlyIleIleMetG 1251
:|||||:::|||||:::|||||:::|||||
3480 GAAGTCAAAATCGCCGCCCTTACCTACCTTATGTCATCATTTTGG 3529
1251 luthrValProGlnLeuTyrAspPheThrGluThrHisAsnGlnArgGly 1267
:::|||||:::|||||:::|||||
3530 ATGCTTTGCCACAGCTCTGTGACTTTACAGTTGCAGATACTCGCAGATAC 3579
1268 ArgProIleCysIleAlaThrAspAspTyrGluSerGluSerGlySerMe 1284
||| |:::|||||:::|||||:::|||||
3580 CGC.....ACCAGTGGCTGGATGAAGAACAAAGAGGACCGGTGC 3620
1284 tIleSerGlnThrValAlaMetAlaIleAlaGlyThrSerValProGlnL 1301
|:::|||||:::|||||:::|||||:::|||||
3621 CATTAACCAAGATGTGGCTGTGCCATAGCAGGGAATTAATTCAATTGGA 3670
1301 euThrArgProGlySerPheLeuLeuThrSerThrSerGlyArgGlnHis 1317
||| |:::|||||:::|||||:::|||||
3671 AAACA.....AGTGAATAGTCTGTCTCTTGCCCTATAAGCAGTAC 3714
1318 ThrThrPheSerAlaGluSerSerArgSerLeuLeuIleCysLeuLeuTr 1334
:::|||||:::|||||:::|||||:::|||||
3715 AACATGCTGAACGCGGACACTACTCGCAACCTCATGATCTGCTTCCTG 3764
1334 pValLeuLysAsnAlaAspGluThrValLeuGlnLysTrpPheThrAspL 1351
|:::|||||:::|||||:::|||||:::|||||
3765 GATCATGAATAAATGCTGATCAGACCTCATTTAGGAAGTGATGCTGACC 3814
1351 euSerValLeuGlnLeuAsnArgLeuLeuAspLeuLeuTyrLeuCysVal 1367
||| |:::|||||:::|||||:::|||||
3815 TGCCATCAACGCGAGCTCAACAGGATTTTATGATCTACTTTTCATCTGTGTG 3864

5641 ATCAAGACCAGGATCAGCGTCATCCAGAAGGAGAGTTGTTTGACACC 5690
1982 oilegluValaIailegluAspmetGlnlysThrglnGluLeuAlaP 1999
|||||
5691 GATTGAAGTTGCCATTGAAGACATGAGAAGAACCCCTGCAGTTAGCAG 5740
1999 heaIatHrHisGlnsPProAlaaspProLysNetLeuGlnMetValLeu 2015
||| :|||
5741 TTGCCATTAAACGAGAGCCGCGCTGATGCAAGATGCTTCAGATGTCCTG 5790
2016 GlnGlySerValGlyThrThrValAsnGlnGlyProLeuGluValAlaG1 2032
|||||
5791 CAAGGCTCTGTGGAGCTACTGTAATCAGGAGACCACTGGAAGTAGCCCA 5840
2032 nValPheLeuSerGluIleProSerAspProLysLeuPheArGHisHisA 2049
|||||
5841 AGTGTTTTGGCTGAATTCCTGCTGATCCAAACTCTATCGACATCACa 5890
2049 snLysLeuArGLeuCysPheLysAspPheThrLysArGcysGluAspAla 2065
|||||
5891 ACAAGTTGAGGTTATGCTTTTAAGGAATTCATCATGAGATGTGGTGAAGCT 5940
2066 LeuArGlysAsnLysSerLeuIleGlyProValGlnLysGluTyrgLnaR 2082
:::|
5941 GTAGAGAAAAACAAGCGCTCATCACGCGACAGACAGGAATATACAGCA 5990
2082 gGluLeuGlyLys 2086
:|||||
5991 GGAAGCTCAAAAAG 6003

seq_name: gb_pr:AB037816

seq_documentation_block:

LOCUS AB037816 4886 bp mRNA linear PRI 14-MAR-2000
DEFINITION Homo sapiens mRNA for KIAA1395 protein, partial cds.
ACCESSION AB037816
VERSION AB037816.1 GI:7243170

KEYWORDS

SOURCE Homo sapiens brain cDNA to mRNA, clone_lib:PBluescriptII SK plus
clone:hj07927.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS Nagase,T., Kikuno,R., Ishikawa,K.I., Hirose,M. and Ohara,O.
TITLE Prediction of the coding sequences of unidentified human genes.
XVI. The complete sequences of 150 new cDNA clones from brain which
code for large proteins in vitro

JOURNAL DNA Res. 7 (1), 65-73 (2000)

MEDLINE 20181126

REFERENCE 2 (bases 1 to 4886)

AUTHORS Ohara,O., Nagase,T. and Kikuno,R.

TITLE Direct Submission
JOURNAL Submitted (31-JAN-2000) Osamu Ohara, Kazusa DNA Research Institute,
Laboratory of DNA Technology; 1532-3 Yana, Kisarazu, Chiba
292-0812, Japan (E-mail:cdnainfo@kazusa.or.jp,
URL:http://www.kazusa.or.jp/huge/, Tel:+81-438-52-3913,
Fax:+81-438-52-3914)

FEATURES

source Location/Qualifiers

1..4886

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="hj07927"

/tissue_type="brain"

/clone_lib="PBluescriptII SK plus"

gene 1..4884

CDS /gene="KIAA1395"

<1..>4884

/gene="KIAA1395"

/note="Start codon is not identified."

/codon_start=1

/product="KIAA1395 protein"

/protein_id="BAA92633.1"

/db_xref="GI:7243171"
/translation="DARTMAASERRAFANKINRTVAAEVRKOVSRERSGSPHSRCS
SSLGVLTEVEPLDEFVLLSRPPDAEPPLRDVLVEFPADDELLLPRECRTEBP
IPKDEKIDAOVRAVEMYIEDWYIVHRYOYISAASPVPTPTQEROKGLPROVEQ
DASGDERSGPEDSNDSSRGSGSPEDTPRSSGASSIFDLRNIAADSLPSLIERAED
VDRNETLRQHRPALLTLYPAPDEDEAVERCSRPEPRERHQRILVKLSLKEI
EIEPIFGLADVDREKKISENFYFDLNSDMKGLRAHGTHPAISTLASAIFSVT
YSPDIFLVTKLEKVLQGGDISCECPYMLKEDVTAKNKEKLEKRLAEQFCRLG
RYRMPFAWTVVHLANIVSSAGQSDSEGERPAWTDRTRRRGQDRASSGDDACSF
SGFRPATLVYTNFQKQAEERLSDLEKFLADMRRPSSLRLRLPVTALQKIDISAP
ENPHFCLSPDLHIKPYDPGRGRKEILEFPAREVYAPHTSYNLLVYPHSLNFS
RQGSVRNLAVRYQYMTGEDPSQALPYFGKSSCSEFTREATFPVYHNKSPFEYEEK
LHPACVTENHLLFTFYHVSQCPREGTALETPVGFTWIPLOHRLRTGPFCLPVS
DQPPSYSVLTPDVALPGMRWVDGKGVSEVELTAVSSVHPDDPYLDKEFTLVHYLEE
GAFPERLKDVTYSEGNVEQELRASLALRLASPEPLVAFSHVLDKLVRLVIRPPIIS
GQIVNLGRGAEAMAHVSVLVHRSLEAAQDARGHCPOLAAIYHYAFRLPGTEPSLPDG
APVTYQAATLARGSGRPASLYLARKSISSNPDLAVAPGSYDDEVSRLASKLHE
ELALQWVSSSAVREAILQHAWFEFQLMVKSMAHLHLGRLDTPRKLREPGREFLDI
TALVGSVGLEVITRVHKDVELAEHLNASIAFEISDLISLVDRGEVSVRAHYKQVAT
RLQSSPNPAALLTRMEFTRIICSHENVTLNLPCCPLSPASPSPSVSTTSQSTF
SSQAPDPKVTSMFELSGPFROQHFLAGILLTELALALEPEAGFLLHKKAISAVHSL
LCGHDTDPRYAEATVKARVAELYLPLISTARDTLPRLHDFAEGBGORSRLASMLDSDT
EGEGDIAGTINDSVAMAIAAGPLAPGSRASISQPPPTASRAGCALSAESSRTLACVL
WLKNTERPALQRMATDITLPLQGRLLDLILCLIAFAEYKKAERINSLTEFKSLD
MKARLEAAILGTIGARQEMVRRSRERSPENVRWRKSVTHMKQTSRDVDTKDEM
EHEALVEGNLATEASLVLDITLLETIVQTYMLSEARESVLGAVLKVYLSLSAQSALE
LOHGLATORALYSKEPELFEEDTELADLCRLLRHCGSRISTIRTHASASLYLMR
QNFIEIGHNEARVKMQVMTMSLSSLVGTQNFSEHLRSLKITLYAEEDMGLRSTFA
EVOVDLMFNLHMLITDVTVMKEHQEDPEMLIDLMYRIARGYGSPDLRLTWLQNMAGK
HAELGNHAEAAQCMVHAA"

BASE COUNT 941 a 1605 c 1425 g 915 t
ORIGIN

alignment_scores: Quality: 5472.50 Length: 1711
Ratio: 3.835 Gaps: 11
Percent Similarity: 83.402 Percent Identity: 61.718

alignment_block:

US-09-737-246-2 x AB037816 ..

Align seg 1/1 to: AB037816 from: 1 to: 4886

2 AlagluArGAlaPheAlaGlnLysIleSerArgThrValAlaAlaG1 18
:::|
22 TCCGAGCGCGCGCTTCGCGCACAGATCAACAGACGGTGGCCGACAGA 71
18 uValArGlyGlnIleSerGlyGlnTySerGlySerProGlnLeuLeu1 35
|||||
72 GGTGCGGAGACGAGGTGTCCCGGACGACGAGTGGCTCCCCCACTCCAGCA 121
35 ySAsnLeuAsnIleValGlyAsnIleSerHisThrThrValProLeu 51
::|
122 GG.....CGCTGCAGCAGCTCCCTGGGGGTCCCACTG 153
52 ThrGluAlaValAspProValAspLeuGluAspTyrrLeuIleThrHisPr 68
|||||
154 ACTGAAGTTGTGAGCCCGCTGACTTTGAGGATGTACTTCTGAGCCGGGCC 203
68 oLeuAlaValAspSerGlyProLeuArGAspLeuIleGluPheProProA 85
| :|
204 ACCAGATGCTGAGCCCGGCCCTCAGGACCTGTAGAAATGCCACGAGCCC 253
85 sPAspIleGluValValTyrrSerProArGAspCysArgThrLeuValSer 101
|||||
254 ATGACTTGAGCTGCTGTCAGACCCCGGGAATGCCGAGCACGAGAGCCC 303
102 AlaValProGluGluSerGluMetAspProHisValArGAspCysIleAr 118
:::|
304 GGGATCCCCAAGATGAACAACTGATGCCACAGGTGAGGGCGCGGTGGA 353
118 gSerTyrrThrGluAspTrpAlaIleValIleArGlySTyrrHisLysLeuG 135

354 GATGTATATGAGGACTGGGTCATGTCACAGAGAAGGTATCAGTACCTGA 403
135 lYThrGlyPheasnProasnThrLeuasplysGlnLysGluArgGlnLys 151
404 GTGCAGCATACAGCCCCGCTCACCACAGACACACAGCGGAGCGACAGAAG 453
152 GlyLeuProLysGlnValPheGluSeraspGluAlaProaspGlyAsnse 168
454 GGCTTCCCCCGCCAGGCTTGTGAGCAGGATGCTTGTGAGCAGAGAGGTC 503
168 rTyrglnaspaspGlnaspaspLeuLysargArgSermetserIleaspa 185
504 CGGCCCTGAGGACTCGAATGACTCCCGCGCTGCGCTCGGCTCCCGCGAAG 553
185 sPThrProargGlySerTrpAlaCysSerIlePheaspLeuLysAsnser 201
554 ACACCCCTCGAAGCAGTGGTGCCCTAGCATCTTCGACCTGAGGAACCTG 603
202 LeuProaspAlaLeuLeuProasnLeuLeuAspArgThrProasnGluGl 218
604 GCAGCTGACTCATGTGTCGCCCTCTGCTAGAGCGGGCGGCCCAAGAAGA 653
218 uIleaspArgGlnAsnaspaspGlnArgLysSerAsnArgHisLysGluL 235
654 TGTGGACCGCGGCAATGAAACCCCTTCGACGGCAGCACCGGCCCGGCC 703
235 euPheAlaLeuHisProserProaspGluGluGluProIleGluArgLeu 251
704 TGCTCACCCCTTACC CGCACCTGACGAGATGAAGCCGTGGAAACGCTGT 753
252 SerValProaspIleProLysGlnHisPheGlyGlnArgLeuValLy 268
754 AGCGGCCAGAGCCACCCCGCGAGCACTTGGACAAAGGATCTTGGTCAA 803
268 sCysLeuSerLeuLysPheGluIleGluIleGluProIlePheAlaSerL 285
804 GTGTGTGTCGCTCAAGTTCGAGATTGAATTGAGCCCATCTTGGGATCT 853
285 euAlaLeuTyraSpValLysGluLysLysIleSerGluAsnPhetYr 301
854 TGGCTGTGATGATGTGCGGGAAGAAAGAAGATCTCGGAGAACTTCTAC 903
302 PheaspLeuAsnSerGluGlnMetLysGlyLeuLeuArgProHisValPr 318
904 TTCGACCTGAACCTCGGACTCCATGAAGGGGCTGCTCGGGCTCATGGCAC 953
318 oProAlaAlaIleThrThrLeuAlaArgSerAlaIlePheSerIleThrt 335
954 CCACCCCTGCCATCTCCACCCCTGGCGCCCTCTGCCATCTTCTGTGACCT 1003
335 yrProserGlnaspValPheLeuValIleLysLeuGluLysValLeuGln 351
1004 ACCCTTCACCTGACATCTTCTGCTCATCAAGTTGGAGAGGTGCTTCAG 1053
352 GlnGlyAspIleGlyGluCysAlaGluProTyrmetiLePheLysGluAl 368
1054 CAAGGGGACATCAGTAGTGTGTGAGCCTTACATGTGTGAAGAAGT 1103
368 aaSpAlaThrLysAsnLysGluLysLeuGluLysLeuLysSerGlnAlaA 385
1104 GGACACAGCCAAAGAACAAAGAGAAGCTAGAGAAGCTGGCGCTGGCGCG 1153
385 spGlnPheCysGlnArgLeuGlyLysTyraArgmetProPheAlaTrpThr 401
1154 AGCAGTTCGACACCGCGCTGGCGCGCTACCGCATGCCCTTCGCTGGACG 1203
402 AlaIleHisLeuMetAsnIleValSerSerAlaGlySerLeuGluArgAs 418
1204 GCCGTGCACCTTGGCCAACATCTGAGCAGCGCTGGGCGAGCTGACCGGGA 1253
418 pSerThrGluValGluIleSerThrGlyGluArgLysGlySerTrpSerg 435

1254 CTCTGAC.....TCGGAGGGCGAGCGCGCCAGCCCTGACAG 1291
435 luArgArgAsnSerSerIleValGlyArgArgSerLeuGluArgThrThr 451
1292 ACCGGCGCGGT.....CGGGGGCCCCAGGACCGGGCGAGT 1326
452 SerGlyaspaspAlaCysAsnLeuThrSerPheArgProAlaThrLeuTh 468
1327 AGTGGGAGCAGCCCTGCAGCTTCTCTGGCTTCCGTCAGCCAGCCTAAC 1376
468 rValThrAsnPhePheLysGlnGluGlyAspArgLeuSeraspGluAspL 485
1377 TGTCAAAACTTCTTTAAGCAGAGGCTGAGCGCACTCAGTGACGAGGACC 1426
485 euTyrlYsPheLeuAlaAspMetArgArgProSerSerValLeuArgArg 501
1427 TCTTCAAGTTCCTGGCTGACATGAGCGCGCGTGTCCCTGCTGGCGCGA 1476
502 LeuArgProIleThrAlaGlnLeuLysIleaspIleSerProAlaProGl 518
1477 CTACGTCCTGTGACTGCCAGCTCAAGATCGACATTCTCCGGCTCTCGA 1526
518 uAsnProHisTyrcysLeuThrProGluLeuLeuGlnValLysLeuTyrr 535
1527 AAATCCCCACTTCTGCTCCTCCCTGAGCTGCTCATATACAGCCCTACC 1576
535 roAspSerArgValArgProThrArgGluIleLeuGluPheProAlaArg 551
1577 CGGACCCCGAGGGCGCGGCCACCAAGAGATTTCTGAGTTCCCGCCCGCC 1626
552 AspValTyraValProasnThrThrTyraArgAsnLeuLeuTyrlIeTyrr 568
1627 GAAGTCTATGCCCCCATACAGCTACAGGAACCTGCTGTACGTGTACCC 1676
568 oGlnSerLeuAsnPheAlaAsnArgGlnGlySerAlaArgAsnIleThrv 585
1677 GCACAGCCTCAACTTCAGCAGCGCGGAGGCTCCGTGCGCAACCTTGTG 1726
585 alLysValGlnPheMetTyrglyGluAspProSerAsnAlaMetProVal 601
1727 TCGAGTGCAGTACATGACAGCGGAGGACCCAGCGAGGCTGTGCCGTC 1776
602 IlePheGlyLysSerSerCysSerGluPheSerLysGluAlaTyrrThrAl 618
1777 ATCTTTGGCAAGTCCAGCTGCAGTGAATTACCCGCGAGGCCCTCACACC 1826
618 aValValTyrrHisAsnArgSerProaspPheHisGluGluIleLysValL 635
1827 GGTGCTACACATAACAAGTCCCCCGAGTCTACGAGGAGTCAAGCTGC 1876
635 ysLeuProAlaThrLeuThrAspHisHisLeuLeuPheThrPheTyrr 651
1877 ATCTTCCAGCCTGCGTGACAGAGAACCATCACCTGCTGTTCACCTTCTAC 1926
652 HisValSerCysGlnGlnLysGlnAsnThrProLeuGluThrProValGl 668
1927 CATGTCAAGTGCAGCGCCCGCGGCGACTGCCCTGGAGACACCGTGGG 1976
668 yTyrrThrTrpIlePrometiLeuGlnAsnGlyArgLeuLysThrGlyGlnP 685
1977 CTTTACTTGATCCCACTGCTGACAGACAGCGCGCTGAGGACCGGCCCT 2026
685 heCysLeuProValSerLeuGluLysProProGlnAlaTyrrSerValLeu 701
2027 TCTGTCTCCAGTGTCTGTGACACAGCGCGCCAGCATACTCCGTGCTC 2076
702 SerProGluValProLeuProGlyMetLysTrpValAspAsnHisLysGl 718
2077 ACACCCGATGTGGCGCTTCCGGGCGATGCCCTGGGTGAGCGGTCAAGAAGG 2126
718 yValPheAsnValGluValValAlaValSerSerIleHisThrGlnAsp 735
2127 CGTGTTCAGTGTGAGCTCACGGCGGTCTCTGTGTGACCCCCCAGGACC 2176

735 roTyrLeuAspLysPhePheAlaLeuValAsnAlaLeuAspGluHisLeu 751
|||||
2177 CCTACCTGGACAATTTTCACCCCTGGTGACAGTCTTGAGAGGAGGCC 2226
752 PheProValArgIleGlyAspMetArgIleMetGluAsnAsnLeuGluAs 768
|||||
2227 TTCCCATTCGCGCTCAAGGACACTGTGCTGAGCGAGGCCAACGTGAGCA 2276
768 ngLuleuLysSerSerIleSerAlaLeuAsnSerSerGlnLeuGluProV 785
:|||||
2277 GGAGCTGGGGCCAGTCTTGACAGCACTGGCGCTGGCCAGCCCGAAACCCC 2326
785 alValArgPheLeuHisLeuLeuLeuAspLysLeuIleLeuValIle 801
::|||
2327 TTGTGGCCTTCTCCACACAGCTGTGGACAAGCTCGTGGCTGTGTCATC 2376
802 ArgProProValIlealagIlglnIleValAsnLeuGlyGlnAlaSerPh 818
|||||
2377 AGGCCCCCGATCATCAGTGGCCAGATTGTGAACCTGGCGCTGAGCCCTT 2426
818 eGluAlaMetAlaSerIleIleAsnArgLeuHisLysAsnLeuGluGlyA 835
|||||
2427 TGAAGCAATGGCCCATGTAGTCAGCCTGTTCACCCGAGCCTGAGGCAG 2476
835 snHisAspGlnHisGlyArgAsnSerLeuLeuAlaSerTyrIleHisTyr 851
::|||
2477 CCCAGGATGCCCGCGGTCACTGCCACACAGCTGGCTGCCCTACGTCCACTAC 2526
852 ValPheArgLeuProAsnThrTyrProAsnSerSerSerProGlyProG 868
::|||||
2527 GCCTTTGGCCTTCTCTGGCACTGAGCCACAGCCTCCCGGATGGGGCCCT. 2574
868 yGlyLeuGlyGlySerValHisTyrAlaThrMetAlaArgSerAlaValA 885
::|||
2575CCAGTGACAGTGAAGGCTGCCACACTGGCCGTGGCTGTGTC 2617
885 rgProAlaSerLeuAsnLeuAsnArgSerArgSerLeuSerAsnSerAsn 901
|||||
2618 GCCCGCAAGCCTTACTGCGCGGCTTCCAAAGACATCAGCAGCAACAAC 2667
902 ProAspIleSerGlyThrProThrSerProAspAspGluValArgSerIl 918
|||||
2668 CCTGACCTGCGCGTGGCCCTGGCTCTGTGATGACGAGGTTTCCCGCAT 2717
918 eIleGlySerLysGlyLeuAspArgSerAsnSerTirValAsnThrGlyG 935
|:|
2718 CCTGGCCAGC..... 2727
935 lProLysAlaAlaProTirPglySerAsnProSerProSerAlaGluSer 951
2727 2727
952 ThrGlnAlaMetAspArgSerCysAsnArgMetSerSerHisThrGluTh 968
2727 2727
968 rSerSerPheLeuGlnThrLeuThrGlyArgLeuProThrLysLysLeuP 985
|||
2728AAGCTGC 2734
985 heHisGluGluLeuAlaLeuGlnTirPValValCysSerGlySerValArg 1001
:|||||
2735 TTCACGAGGAGGCTGGCTGTGAGTGGGTGTCAGCAGCAGTGCCGTACGC 2784
1002 GluSerAlaLeuGlnGlnAlaTirPhePhePheGluLeuMetValLysSe 1018
|:|
2785 GAGGCCATCTCTCCAGCAGCGCTGGTCTTCTTCCAGCTCATGGTGAAAGAG 2834
1018 rMetValHisHisLeuTyrPheAsnAspLysLeuGluAlaProArgLys 1035
|:|
2835 TATGGCGCTGCACCTGTGCTTGGCCAGCGACTAGACACACCCCGCAAGC 2884

1035 erArgPheProGluArgPheMetAspIleAlaAlaLeuValSerThr 1051
|||||
2885 TGGCCTTCCCGGAGCGCTTCTTGAGCAGACATCACTGCTTGGTGGCTCT 2934
1052 IleAlaSerAspIleValSerArgPheGlnLysAspThrGluMetValGl 1068
::|
2935 GTGGCGCTGGAGGTTCATCACCCGTTGTCACAAAGATGTGAGCTGGCCGA 2984
1068 uArgLeuAsnThrSerLeuAlaPhePheLeuAsnAspLeuLeuSerValM 1085
|:|
2985 GCACCTCAACGCCAGCCTGGCTTCTTCCCTCAGTGACCTTCTGTCCCTGG 3034
1085 etAspArgGlyPheValPheSerLeuIleLysSerCysTyrLysGlnVal 1101
:|||||
3035 TGGACCGGGGCTTTGTCTTACGCTGTGCCGGGCCACTACAAAGCAGGTG 3084
1102 SerSerLysLeuTyrSerLeuProAsnProSerValLeuValSerLeuAr 1118
::|
3085 GCCACGCGGCTCCAGTGTGCTCCCTTAATCCAGCAGCCCTGTGACCTGGC 3134
1118 gLeuAspPheLeuArgIleIleLysSerHisGluHisTyrValThrLeuA 1135
|:|
3135 CATGGAATTCACCCGCATCTGTGAGCCACGACGACACTAGTGAACCTCA 3184
1135 snLeuProCysSerLeuLeuThrProProAlaSerProSerProSerVal 1151
|||||
3185 ACCTCCCGCTGGCCCCCTGTACCTCCAGCCTCGCCCTCCCTCTGTG 3234
1152 SerSerAlaThrSerGlnSerSerGlyPheSerThrAsnValGlnAspGl 1168
|:|
3235 TCCTCCACCACTCCACAGAGCTCCACCTTCTCCAGCCAGCCCGGAGCCC 3284
1168 nLysIleAlaAsnMetPheGluLeuSerValProPheArgGlnGlnHisT 1185
|:|
3285 CAAGGTGACACGATGTTGCACTGAGTGGACCATTTCCGGCAGCAGCACT 3334
1185 yrLeuAlaGlyLeuValLeuThrGluLeuAlaValIleLeuAspProAsp 1201
:|||||
3335 TCCTAGCTGGGCTCTCTGTGACGAGAGCTGGCACTGGCCCTCGAACTGAG 3384
1202 AlaGluGlyLeuPheGlyLeuHisLysLysValIleAsnMetValHisAs 1218
|:|
3385 GCTGAAGGGCATTCCTGTGTGACACAGAAGCCATCAGTGTGTGCACAG 3434
1218 nLeuLeuSerSerHisAspSerAspProArgTyrSerAspProGlnIleL 1235
:|||||
3435 CCTGCTATGTGGCCATGACACTGACCCCCGCTACGCCGAGGCCACTGTGA 3484
1235 ysAlaArgValAlaMetLeuTyrLeuProLeuIleGlyIleIleMetGlu 1251
|:|
3485 AGGCTCGTGTGGCCGAGCTGTACTGTGCCACTGTCTTTCGATTGCACGGGAT 3534
1252 ThrValProGlnLeuTyrAspPheThrGluThrHisAsnGlnArgGlyAr 1268
|:|
3535 ACCTTGCCACGGCTGCATGACTTGTGAGGGGCCAGGTCAGCGGTCAAG 3584
1268 gProIleCysIleAlaThrAspAspTyrGluSerGluSer.....Glys 1283
|:|
3585 ACTGGCTCAATGCTTGACTCAGACACAGAAGCGAAGGCGACATTGCGG 3634
1283 erMetIleSerGlnThrValAlaMetAlaIleAlaGlyThrSerValPro 1299
:|
3635 GTACCATCAACCCCTCTGTGGCCATGGCCATTGCTGTGGCCCCCTA... 3681
1300 GlnLeuThrArgProGlySerPheLeuLeuThrSer.....Th 1312
|:|
3682GCCCTGGCTCCCGGGCCAGCATCTCCAGGGGCCACCAAC 3722
1312 rSerGlyArgGlnHisThrThrPheSerAlaGluSerSerArgSerLeuL 1329
|:|
3723 GGCTTCTCGCGCAGGCTGTGCCCCCTCTGTGAGTCAAGCCGACCTTGC 3772
1329 euIleCysLeuLeuTirPValLeuLysAsnAlaAspGluThrValLeuGln 1345

3773	TCGCGGTGTGTGCTGTGGGTGCTGCAAAAAACACCGAGCGCGGCTTCGTGCAG	3822
1346	LysTrpPheThrAspLeuSerValLeuGlnLeuAsnArgLeuLeuAspLe	1362
3823	CGCTGGGCCACTGACCTGACACTGCCCCACAGCTGGGACGCTGTGGACTT	3872
1362	uLeuTyrLeuCysValSerCysPheGluTyrLysGlyLysValPheG	1379
3873	GCTGTACCTTTGCCCTGGCTGCCCTTTGAGTACAAAGGGAAAAAGGCTTTG	3922
1379	luArgMetAsnSerLeuThrPheLysLysSerLysAspMetArgAlaLys	1395
3923	AACGCATCAACAGCCTCACATTCAAAAAATCTGTGATATGAGGCGCGG	3972
1396	LeuGluGluAlaIleLeuGlySerIleGlyAlaArgGlnIleMetValAr	1412
3973	CTAGAGGAAGCCATTCTGGGTACCATCGGAGCTCGACAAGAAATGGTTGCG	4022
1412	gArgSerArgGlyGlnLeuGluArgSerProSerGlySerAlaPheGlys	1429
4023	GCGAAGTCGT.....GAGAGGAGCCCG.....TTTGGGA	4051
1429	erGlnGluAsnLeuArgTyrArgLysAspMetThrHisTyrArgGlnAsn	1445
4052	ATCCAGAGAAGCTGCCGTGGCGGAAGAGCGTCACACACTGGAAACCAACC	4101
1446	ThrGluLysLeuAspLysSerArgAlaGluIleGlnHisGluAlaLeuIl	1462
4102	TCAGACCCGCGTGACACAGACCAGATGAATGGAACACAGAGCCCTTGGT	4151
1462	eAspGlyAsnLeuAlaThrGluAlaAsnLeuIleIleLeuAspThrLeuG	1479
4152	GGAAGGGAACCTGGCAACCGAGCAAGCCTAGTGGTTCTTGACACACTGG	4201
1479	IuIleValValGlnThrValSerValThrGluSerLysGluSerIleLeu	1495
4202	AGATCATCGTGCAGACGGGTGATGCTTTCAGAAAGCCCGGAGAGCTCTTG	4251
1496	GlyGlyValLeuLysValLeuLeuHisSerMetAlaCysAsnGlnSerAl	1512
4252	GGGGCAGTGCTGAAGGTGTGCTGTACAGCCTGGGCAAGTGCACAGAGTGC	4301
1512	aValTyrLeuGlnHisCysPheAlaThrGlnArgAlaLeuValSerLysP	1529
4302	CCTCTTCTTGCAGCATGGCCTGGCCACCCAGAGGGCCCTGTGTCCAACT	4351
1529	heProGluLeuLeuPheGluGluGluThrGluGlnCysAlaAspLeuCys	1545
4352	TCCCGGAGCTGCTGTTCGAGGAGGACACGGAGCTGTGTGCCGACCTGTGC	4401
1546	LeuArgLeuLeuArgHisCysSerSerSerIleGlyThrIleArgSerHi	1562
4402	CTGAGGCTCCTACGACACTGGGCAGCCGATCAGCACCATCCGACGCA	4451
1562	sProSerAlaSerLeuTyrLeuLeuMetArgGlnAsnPheGluIleGlyA	1579
4452	CGCCAGCGCCTGCTGTACTGCTCATGCGACAGAACTTCGAGATCGGCC	4501
1579	sAsnPheAlaArgValLysMetGlnValPrometSerLeuSerSerLeu	1595
4502	ACAACCTTGGCCGCTGAAGATGCAGGTCACCATTGTCTCTTCCTCCCTG	4551
1596	ValGlyThrSerGlnAsnPheAsnGluGluPheLeuArgArgSerLeuLy	1612
4552	GTTGGGAGCAGCAGCAACTTCAGTGAAGAGCACCTGCGAGCTTCACCTCAA	4601
1612	sThrIleLeuThrTyrArgGluGluAspLeuGluLeuArgGluThrThrp	1629
4602	AACCATCTCTACCTATGTCTGAGGAGGACATGGGCTGCGGAGACACACT	4651
1629	heProAspGlnValGlnAspLeuValPheAsnLeuHisMetIleLeuSer	1645

```

4652 TCGCAGAGCAGGTCCAGGACCTGATGTTCAACCTGCACATGATCCTGACG 4701
1646 AspThrValLysMetLysGluHisGlnGluAspProGlnMetLeuIleAs 1662
|||||
4702 GACACGGTGAAGATGAAGGAACACACAGAGACCCCTGAGATGCTCATCGA 4751
1662 pleuMetTyrArgIleAlaLysGlyTyrGlnThrSerProGlu...ArgL 1678
|||||
4752 CCTCATGTACAGAAATTGCCCGGGGCTACACGAGCTCACCGGACCTTCGGC 4801
1678 eUThrTrpLeuGlnAsnMetalagLysHisSerGluArgSerAsnHis 1694
|||||
4802 TGACCTGGTTCGACAGACATGGCCGGGGAAGCACGCGGAGCTGGGCAACAC 4851
1695 AlaGluAlaAlaGlnCysLeuValHisSerAla 1705
|||||
4852 GCCGAGGCCGCCAGTGCATGTTGCACGCCGCC 4884

seq_name: gb_ov:AY016023

seq_documentation_block:
LOCUS AY016023 167215 bp DNA linear VRT 01-FEB-2002
DEFINITION Sphoeroides nephelus alpha globin gene cluster, complete sequence.
ACCESSION AY016023
VERSION AY016023.1 GI:18463963
KEYWORDS
SOURCE
ORGANISM
Sphoeroides nephelus.
Sphoeroides nephelus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Sphoeroides.
1 (bases 1 to 167215)
1 Flint,J., Tufarelli,C., Peden,J., Clark,K., Daniels,R.J.,
Hardison,R., Miller,W., Phillips,S., Tan-Un,K.C., McMorrow,T.,
Frampton,J., Alter,B.P., Frischauf,A.M. and Higgs,D.R.
Comparative genome analysis delimits a chromosomal domain and
identifies key regulatory elements in the alpha globin cluster
Hum. Mol. Genet. 10 (4), 371-382 (2001)
JOURNAL
MEDLINE
PUBMED
21096913
11157800
REFERENCE
2 (bases 1 to 167215)
1 Flint,J., Tufarelli,C., Peden,J., Clark,K., Daniels,R.J.,
Hardison,R., Miller,W., Phillips,S., Tan-Un,K.C., McMorrow,T.,
Frampton,J., Alter,B.P., Frischauf,A.M. and Higgs,D.R.
Direct Submission
Submitted (06-DEC-2000) MRC Molecular Haematology Unit, Institute
of Molecular Medicine, John Radcliffe Hospital, Oxford, Oxon OX3
9DS, UK
FEATURES
source
1.167215
/organism="Sphoeroides nephelus"
/db_xref="taxon:39110"
1510.1530
/note="simple repeat"
/rpt_type=tandem
/rpt_unit=c
2315.2350
/note="C-rich; low complexity"
2643.2663
/note="simple repeat"
/rpt_type=tandem
/rpt_unit=ca
3620.3676
/note="simple repeat"
/rpt_type=tandem
/rpt_unit=tg
4231.4284
/note="simple repeat"
/rpt_type=tandem
/rpt_unit=ca
4638.4678
/note="simple repeat"

```

repeat__region /rpt_type=tandem
/rpt_unit=tg
5566. .5600
/note="simple repeat"
/rpt_type=tandem
/rpt_unit=ca
5840. .5861
/note="AT-rich; low complexity"
8334. .8354
/note="AT-rich; low complexity"
9950. .9970
/note="AT-rich; low complexity"
10777. .10802
/note="AT-rich; low complexity"
join(<10872. .11080,11797. .>11973)
/gene="unkempt"
/product="unkempt"
<10872. .>11973
/gene="unkempt"
join(<10872. .11080,11797. .>11973)
/gene="unkempt"
/codon_start=1
/product="unkempt"
/protein_id="AAL73045.1"
/db_xref="GI:18463964"
/translation="YLKEFRTEQCPLFLQHKCTOHRPFTCFNNHFLNQRRRPIRRRD
GTFNYSPPVYCTKYDETGTGLCPDGDCCPYLHRTTGTERTKYHLRYKKTGTCIHETDAR
GHCVKNGLHCAPFAHGPHDLRPVYDI"
17726. .17763
/note="A-rich; low complexlty"
19346. .19371
/note="simple repeat"
/rpt_type=tandem
/rpt_unit=c
20068. .20121
/note="AT-rich; low complexity"
complement(<21113. .>22747)
/gene="IGFALS"
/product="IGFALS"
complement(<21113. .>22747)
/gene="IGFALS"
complement(21113. .>22747)
/gene="IGFALS"
/codon_start=1
/product="IGFALS"
/protein_id="AAL73046.1"
/db_xref="GI:18463965"
/translation="ACARGCTCQHDDYSLLELVYCSAQNLTQVPPDMPSTHSLWLDV
NLFASLPASFKDVLHLEFLINLSGOLVTLDPQALGRLSLAHILERNRIRVLPGAI
EQNTPKIASLSLHNNQLSRIDRLFAGLSHMLNLGRNSIAVLDETVEFQDLQKKEI
ILVGNRLAYLOPOLFQNLVELKEIDLSGNHLKVIKANVFVKLTKQKLYLAONOIMTV
VPRAFVGMSLRWLDLTNNRLTSLHEDTFMGLYLLHVLRVSNNSIGGIKPRTFRDLQY
LEELRLSHNRIRVYIGEKEVEELGRLEVELEHNRIOEAKVGSFTGISHVAVINLSGSC
FHTLPDQVFRGISKLHSLHLDRGCLTRVTTQAFSGLSLRRLFLQHNNISVEROSFV
DLVGLGLDLSEFNKLEVLTGQTFSGLKNELEYLLSNDCRHFLQNGSRLLLPRLRYLD
LRANELTSIVPEFAEGLEKLLLSGNRMRCDCGALPLRNYSLRNALVYPRQVETHAEGE
EPDTTITYNNITCHSPRLTGGDLRDVSGSVFQSC"
24332. .24356
/note="simple repeat"
/rpt_type=tandem
/rpt_unit=gga
24525. .24554
/note="GC-rich; low complexity"
24862. .24895
/note="AT-rich; low complexity"
26315. .26340
/note="simple repeat"
/rpt_type=tandem
/rpt_unit=ca
26915. .26951
/note="simple repeat"
/rpt_type=tandem
/rpt_unit=cgg

repeat__region 27240. .27272
/note="simple repeat"
/rpt_type=tandem
/rpt_unit=tg
27487. .27507
/note="AT-rich; low complexity"
28056. .28080
/note="simple repeat"
/rpt_type=tandem
/rpt_unit=tg
28159. .28189
/note="simple repeat"
/rpt_type=tandem
/rpt_unit=ccg
30084. .30122
/note="simple repeat"
/rpt_type=tandem
/rpt_unit=cag
30217. .30260
/note="simple repeat"
/rpt_type=tandem
/rpt_unit=tg
31181. .31214
/note="simple repeat"
/rpt_type=tandem
/rpt_unit=ga
33149. .33186
/note="simple repeat"
/rpt_type=tandem
/rpt_unit=tg
33186. .33241
/note="simple repeat"
/rpt_type=tandem
/rpt_unit=ga
33298. .33379
/note="simple repeat"
/rpt_type=tandem
/rpt_unit=tg
33966. .34015
/note="T-rich; low complexity"
35457. .35500
/note="simple repeat"
/rpt_type=tandem
/rpt_unit=ca
35965. .35993
/note="simple repeat"
/rpt_type=tandem
/rpt_unit=ttc
37605. .37637
/note="simple repeat"
/rpt_type=tandem
/rpt_unit=tg
37782. .37804
/note="simple repeat"
/rpt_type=tandem
/rpt_unit=ca
40208. .40383
/note="simple repeat"
/rpt_type=tandem
/rpt_unit=tcca
40386. .40562
/note="simple repeat"
/rpt_type=tandem
/rpt_unit=tcta
41620. .41645
/note="simple repeat"
/rpt_type=tandem
/rpt_unit=ca
41968. .41994
/note="simple repeat"
/rpt_type=tandem
/rpt_unit=tgg
44729. .44773

/note="simple repeat"
/rpt_type=tandem

alignment_scores:

Quality: 4986.50 Length: 5470
Ratio: 2.677 Gaps: 90
Percent Similarity: 34.059 Percent Identity: 26.362

alignment_block:

US-09-737-246-2 x AY016023/rev ..

Align seg 1/1 to reverse of: AY016023 from: 1 to: 167215

```

49 ValProLeuThrGluAlaValAspProValAspLeuGluAspTyrLeuI1 65
   :::::::::::::::::::::
134857 TTGCCCCGTGACAGAGGTGGTGAGCCGGTGATTTTGAGGATTAATGTGAG 134808

65 eThrHisProLeuAlaValAspSerGlyProLeuArgAspLeuI1GluP 82
   :::::
134807 CAGCCATGCTCCTGGGGTGAGACCTGGTCCCTCAGAGAGCTTCTGAGT 134758

82 heProProAspAspIleGluValValTyrSerProArgAspCysArgThr 98
   |||||||
134757 TCCCCCAGACGACCTGGAGCTCTCTCCACCTGGACAAGAGTGCACAACC 134708

99 LeuAlaSerAlaValProGluGluSerGlu..... 108
   |||
134707 CTGGAGCCCCCCTGCTGCTGATGAAGAGAGTAAGTTTGTCCCTGTGCC 134658

108 ..... 108

134657 AACACAATGCTCAACAACACTCTGTCTTCGCTCTTTTAACAACAACACA 134608

108 ..... 108

134607 CACACACACACATACACACAGATGATGATTACATATTGGAGTGGAGG 134558

108 ..... 108

134557 TGAACATGACAGACAAGGCTCTTCGAAGTGGCTGAATAACAACAAAC 134508

108 ..... 108

134507 ACAGTTACTGTATTAAAGCATTTTCTGCTCAATTATTCAATGTGTTGTC 134458

109 .....MetaspProHisValArgAspCysIleArgSerTyrThr 121
   :::::
134457 TGTGTAGCTCACTGGATCCAGAGTGCAGACACCTTGTCACTCTACACA 134408

122 GluAspTrpAlaIleValIleArgLys..... 130
   :::::
134407 GATGACTGGCTCATCATTTCAAGAAAGTAAGTGCTACATGTGAGGCAGA 134358

130 ..... 130

134357 CCGTTAAATGTGTTTCCCTCTGCAGAGACCCCTTTTATCTGTGATTG 134308

131 .....TyrHisLysLeuGly 135
   ||:::
134307 CTTCGTTATTAACTCGGCTCTCTTTTATGTAGATACCAAGCATTAACAGC 134258

136 ThrGlyPheAsnProAsnThrLeuAspLysGlnLysGluArgGlnLysG1 152
   |||
134257 ACCACTTTCACGCCCTCACAGCTCAGAGAGACAGAGAGAGAGAGCGAGG 134208

152 yLeuProLysGlnValPheGluSerAspGluAlaProAspGlyAsnSerT 169
   |||
134207 GCTGGTCAAAACAGACCTTTGAGCTGATGTGGCG.....GCCACTGAGC 134164

169 yRGlNAspAsp..... 172
   :::::
134163 ACCAGGACGACCAAGTAGTAATGGTCACTGTGATCGAGTCTGATCCGAGT 134114
```

```

172 ..... 172

134113 CCCATCAGTGTGAGATTATTGACCACACCCGTATTAAATATATCGATAATGG 134064

172 ..... 172

134063 TGCCAGGAAGAAAAGCAATTCAGAGTTTTTTTTTTTCTTTTAATATGAT 134014

172 ..... 172

134013 AGCTGGACAAGACACCAACCACCTGAAAAAGCTGTGTCTCGATGG 133964

172 ..... 172

133963 GAAGATGAAAAAAGACGAGTACTACAGAAATTGATATACCTGTCATTA 133914

172 ..... 172

133913 TCTCCACATTCCCCATATTATCATATTTAAATAGTAATAAGTTAATGGGGT 133864

172 ..... 172

133863 CTGTATCATATTACTCTATAAAACATCTCTGCAACCACTTAAGGGGGC 133814

172 ..... 172

133813 ATGCTCTCCTGGAAGAAATAAAACACAATGATCTCATTCCTTGTCTCA 133764

173 .....GlnAspAspLeuLysArgSerMetSerIleA 184
   |||||
133763 CCCTCCTTACGATTCACAGAGATGACGCAAAAGCGGCTCAGTCAGCCTCG 133714

184 spAspThrProArgGlySerTrpAlaCysSerIlePheAspLeuLysAsn 200
   :::::
133713 AAGAAACTCCTCGAGGAGCTGGGCTCCAGCAATCTTTGACCTGAAGAAC 133664

201 SerLeuProAspAlaLeuLeuProAsnLeuLeuAspArgThrProAsnG1 217
   |||
133663 TCCACCCACAGATGCTCTCCCTGCCGACTGTCTTAGAACGACGCGCTGGA 133614

217 uGluIleAspArgGlnAsnAspAspGlnArgLysSerAsnArgHisLysG 234
   :::::
133613 GGACATGGACCAACCGCATACCGAGGCGGCCCAAGGGCGCCACAGTG 133564

234 IuLeuPheAlaLeuHisProSerProAspGlu..... 244
   :::::
133563 ACCTGCTGGGCTTTACCCACCCCTGATGAGGTTTCTCAAAAGCATTTT 133514

244 ..... 244

133513 GTGATCTTGAATAATGCAATTTTTTAATAACACTGCACATCAGCAAAAT 133464

244 ..... 244

133463 TCTAACAATATTTCTCAGCAAAACACACAATTTAATCAAAAATTAATTAAT 133414

244 ..... 244

133413 ATTGATAGATTGTCACTGGAATATATGCCGGTCTGACAGATGGAATGTGT 133364

244 ..... 244

133363 GATTTAAGCCTTGTGTTATTTTGAATAGCTTACATTGACCTATGAGACA 133314

244 ..... 244

133313 TTTATGCCACTGTTTAAAAATGCATGACACTTGAATGTTTGTGATGT 133264

244 ..... 244

133263 AAGTGATTTTGTCAAGCTCTGGGTAAAGTAATGTCTTTGTCTCAAAAG 133214

244 ..... 244
```


133213 TGCAGCCATTACAAGAGATGGTTTCTGATTTGTCCTTCCCTTATACA 133164
244 244
133163 ACGTCTGTGTGCACATGTTCACTTCTCTGGCTCCACGTTGCCATGACAT 133114
245GluGluPro 247
133113 CAGGTAAATATTTGTTCTTGGCTGTGCCCCGCGTGACAGACGGGCT 133064
248 IleGluArgLeuSerValProAspIleProLysGluHisPheGlyGlnAr 264
133063 GTCGAAAGATGCTCCACTGTGTAAGTGCCCAAGGAGCACTGTGCCAGAG 133014
264 gLeuLeuValLysCysLeuSerIeu..... 272
133013 GATTATGGTGAAGTGTCTGTCTGAAGTGAGTAGTAAATAATTAGAT 132964
272 272
132963 TTATGTCATTTTAATGTACAGAGCCCATATTTGTTAACAATTCTTCATA 132914
273Lys..PheGluIleGluIleGluProIlePheAlaSerLeuAla 286
132913 CGCTTTTCAGGTTTGAATAATAGAACCGGATATTTGGAATCCTTGCC 132864
287 LeuTyrAspValLysGluLysLys..... 295
132863 CTGTATGACATCAAGGAAAGAAAAAGTATTTTTCCTCATTTAGATG 132814
295 295
132813 ACTTGCACTTACGTAATGCTGAGCTCATGCTCCGCTAATTCATTCCCT 132764
295 295
132763 TGTCTTAAAGGTCTAAATGGCTCTACAGTTTAGATCACCGAGTCAAAAC 132714
295 295
132713 TGAATTAAATTCCCTGACTTGAATTTCCATGATCGTCATCTCTCCTCC 132664
296IleSerGluAsnPheTyrPheAspL 304
132663 TTCCCTGCTTATCTTTCCTGCACAGATCTCAGAGAATTCTACTTCGACC 132614
304 euAsnSerGluGlnMetLysGlyLeuLeuArgProHisValProProAla 320
132613 TGAACTCAGATCAGATGAAGGGCTACTCAAAACACACACGCCCTCACACA 132564
321 AlaIleThrThrLeuAlaArgSerAlaIlePheSerIleThrTyrProse 337
132563 GCCATTTCACCCCTGGCCCGCTTCAGCTATTTTCTCATCACGTACCCCTC 132514
337 rGlnAspValPheLeuValIle..Lys..... 345
132513 TGCCGATATCTTCTTGGTCATTAAAGGTACTGAAGCTGCCCACTTCACCTGT 132464
345 345
132463 ACAATCTGGTTAATGTAAAGCATAAATAGATCAGGTGTGTTGTGCAG 132414
345 345
132413 ACATGATCTAGTGTGCTAGATATACAATGCAATTAATTGTTTGGCGC 132364
346LeuGluLysValLeuGlnGln 352
132363 TAGTCTTAACATGTTTGCATGCATTTCAAGCTGGAAGGCTCCTTCAACAA 132314
353 GlyAspIleGlyGluCysAlaGluProTyrMetIlePheLysGluAlaAs 369
|||||:::|||||:::|||||:::|||||:::|||||:::||

132313 GGAGATATCAGTGAGTGCTGTGAGCCCTACATGGTCATGAAAGAGTCTGA 132264
369 PalAThr..Lys..... 372
132263 TTCTCAAGGTAATGCAGAAAAAATGTACTGCTGTGATGTCACATAACA 132214
372 372
132213 GCCACCATGTAACCTTGACCTGTGAGTCAGAAACATAATTCAATTCTGCTC 132164
372 372
132163 CCAAAAGTTCATTTTGTCTGTGCTGGAACCTTCTAATCTTCTCGACGCC 132114
372 372
132113 TACTCTGAGGAGAGAGACAGTCGACACATTTCCGCTGTAATAAAGAGAA 132064
372 372
132063 CAAGTCTGTTGATTAATAAGAGTGAATAATTACCAGCTCTGTCTGGGTGT 132014
372 372
132013 GTGCGATGTTAGTTTCATTTACAGCTTTTGTGATTTACAGTTTCAACAAA 131964
373AsnLysGluLysLeuGluLysLeuLysSerGln 383
131963 TCACGTTCTCTGTTTCAGCACAAAGAGAGAGCTGGAGAAAGCTGCGTCTGCAG 131914
384 AlaAspGlnPheCysGlnArgLeuGlyLysTyrArgMetProPheAlaTr 400
131913 GCAGAGCAGTCATGCACAACCGCCTTGACGCTTCCGATGCCCTTTGCCCTG 131864
400 pThrAlaIleHisLeuMetAsnIleValSerSerAlaGlySerLeuGluA 417
131863 GACAGCCATCCACCTGCTGTTACATTTGTACAGCAGTATGGGAGCTTGTGATC 131814
417 rGAspSerThrGluValGluIleSerThr..... 426
131813 GGTCAAGACCCGGACTTTGATTTCCGGTACCGTGATTCAGCCBACCTGTATAA 131764
426 426
131763 ATGCTAATGGCCCTTTTACTGTTTCATGCTGTGTAATAATGTCCTTATA 131714
427GlyGluArgLys.....GlySerTrpSerGluArgArgAsn 439
131713 TATGCTAACAGAGAAAAAAGGCCATGGGACATGGAATGAGCGAAAG... 131668
439 erSerIleValGlyArgArgSerLeuGluArgThrThrSerGlyAspAsp 455
131667AAAAAGGGGCTTGAGCGGCATGAGTATTTGCGAGCGAC 131632
456 AlaCysAsnLeuThrSerPheArgProAlaThrLeuThrValThrAsnPh 472
131631 ATGTGCAATTTGGCCACCCTTCCGTCAGACAACATTGACCGCTCACCAACTT 131582
472 ePheLysGln..... 475
131581 CTTCAAGCAGGTCTGACTGTCTGTCAATAGTAAAGCTTGAGTGTGTGCCCT 131532
475 475
131531 ATATTGATTTGTTGTCTATCATCATTAACATACAAAACAATGATGC 131482
475 475
131481 ATTATCTTAAAGACACCACTAGACTGTGATCAAGTAATGTATTGCAACTC 131432
476Glu..G1 477
131431 AAAGTTATAGGGAATAAAGTTAAATTTCCTTTGTTTTCATGAGAGAGGG 131382

477 yAspargLeuSeraspGluspleuTyrllysPheLeuAlaspMetArga 494
|||||
131381 GGACCGACTAAGTGACGAGGACCCTTTATAAGTTCTCGCAGATATGCCGA 131332
494 rgProSerSerValLeuArgArgLeuArgProIleThr..... 506
|||||
131331 GACCGTCTTCTGTCTCGCAGCAGTGGCCCTGTCAACAGGTAAGCATTAA 131282
506 506
131281 GACTTACTTGTACAAAGAGTAATTGTTCTAACATGTGATGGTGTTTGTC 131232
506 506
131231 CTCTCCTAACTTAAGAATGCACGTTAAAGGTGTTAATCCAATTTTCT 131182
507AlaGlnLeuIleasp 512
131181 TCTTTGTGTCTGTGATTATTTCGACCTGCACGCCAGTTGAAGATTGAT 131132
513 IleSerProAlaProGluAsnProHisTyrCysLeuThrProGluLeu 529
|||||
131131 ATCTCTCAGCTCCGGACTCGCCCTCACTACTGTCTGTACCACTGCT 131082
529 uGlnValLysLeuTyrProAspSerArgValArgProThrArgGluIle 546
|:::|
131081 TCATGTAAACCCTACCCAGACCACGGGTTCGCCCAACCAAGAAAGTGC 131032
546 euGluPheProAlaArgAspValTyrValProAsnThrThrTyr..... 560
|||||
131031 TGGAGTCCCTGCTGCTGTTATGTGTACGCCACATACCTCCTATAAGTGTA 130982
560 560
130981 GAGCCCCCAACTCACAGAATCATTTATTTCTTCTCTGTCAATATCAA 130932
560 560
130931 TTAATTATTGCTGCCTTATTGCTTATAATTCTGATAATATTAAAGAG 130882
560 560
130881 GATAAAACACAGTAATACCTAATGAAAAACAATATTTTAAATACATTCT 130832
560 560
130831 AAATAGAAATAACATTCTCAGTAATTGTCATGTTTAGTCAAGTTTAAATA 130782
560 560
130781 ATGTAACAATTGTTTCAGTTATAAAGCTAAGTTCACAAAGTACAGCACAA 130732
560 560
130731 ACTCCCATCATACAGGTCAAAAGATAAAGAACCTGTAGTAAGGTGATA 130682
560 560
130681 GTCAAGTTCAGCTTACGTACATGTACAGGAGAGAAAGAGAGTGCATG 130632
560 560
130631 TTGCAGTAAAAAAACCCTGTGGGGAATCTTAAATGCCAATGCTCTGC 130582
560 560
130581 CTTCGCTCTTCAGATGCAAAAGAAACAAGATACAGCATTAGAGGAT 130532
560 560
130531 ATGAATAAACAACGGGGGCTTTTAAACAATAATCTTACCGTTTCTGTAG 130482

560 560
130481 TTTAGTGAGCTTTTCACTCTGACTCTTGATGCTCTCCTTTTGAACGTGAA 130432
560 560
130431 TCAGTTCTGTAGTCTCTCAAAACTTCAGTGTGCTCCTTCGCTAGTC 130382
560 560
130381 TAGAGCGGGAGAACAGTTTCCATATATCATTTTGTGTTGTATGCTG 130332
560 560
130331 TATTGCTCTTAGTATTAATTACAGAGGAATTCTGTACCTGCAGCTCTTT 130282
560 560
130281 GATGCTGCTGGGTACATACTTATTGAGATGTGCTCCAGCCTCTTTA 130232
560 560
130231 CATTTGCTGAGCTCCACTTCTTCCGTCTCCAGGTTGGCCAGCCTGCT 130182
560 560
130181 TTCGTACCTCGGTCTCAGCCTGTGCTGGGCCATCTGGGTCTGTGAAA 130132
560 560
130131 GAGCATCTGTCCAGTGTATTCTCTCCTACAGGACAATCTCACGGGAGCTG 130082
560 560
130081 GCAGCCGCTGACCCCCGTATAACATGTGGCTTCTAATGACACACTCTC 130032
560 560
130031 AGATGAAGGCTCTGCTATATCTGCAAAACATCATGAAAACCAATGTGAGA 129982
560 560
129981 CCAAAATCACCAAAAGTCTTAGAAAAACAATCTTATCCCCAAAGGACAG 129932
560 560
129931 TGCTTTATTCAGATTTTGGAGATATGACCCGTGTGTTTCTGCCCCCTCTG 129882
560 560
129881 CATGTTTTCGCTGCGAGTTTGATTACAGTGTGTCTCTCTCTCTCT 129832
560 560
129831 AGCAGCACAGCCCTCAGTAGCCAGACACATGTTGTTGACATATTCT 129782
560 560
129781 GCCATCACTTACCTGAGCAACATTAATACTGCTTTAGCCGTGTTTTTCA 129732
560 560
129731 TCTCTGAGCTTCCCTCAGCATGTCTCCCGAGCCTTGAAGTATCCTTTCA 129682
560 560
129681 TGGTTCTTCAAAGTCTGGCTTATTCTGTTTATCTTGGCCTCGGTGTTTC 129632
560 560
129631 AGAATGTATTGAGGGAGTCCCGAACTGTATGAATCCAACGTGAGCA 129582
560 560

129581 CGTTGACCTCGTCCTGTGGTGAACAATTGTTTTTCCTGCTGGTCTTCCTC 129532
560
129531 ACTGACCTGCACGGTTGCTCTTAAACCCCGACGCAACCGATCAGACA 129482
560
129481 GAGACTCCAGATCAGATTCATAGTATATTCCCTCAGTGTAGTCCAGCTG 129432
560
129431 TTTCAGAGTCTCGGTGTTTCTTCTTCAGATTGTAGTCAGAAAAGTTAGT 129382
560
129381 TAACTATTCGGTAAATAGAGTCCCATCTGTTTAAAGATCCAGTGTGTGT 129332
560
129331 GGGTCTGTGTGAATTGGAGGTTAGTCAGGTGCATCTTATATTGGATGTG 129282
560
129281 TATATGCATAGTAAATATTTAGCTCAGCTGTGATGTTCACTCAGTACCCT 129232
560
129231 ACCTCTTTGACCTGCCTTAATGTTTAAACCCCTGAGATACTGATCTTGA 129182
560
129181 CATTTAGGTCAACAACACATGTGACACATTTACTGTATATTTTGCTTTTGA 129132
560
129131 CACAGATCTTGGTGTGTTTTCATGCCACACTGCTGTGTAAATGGTGT 129082
560
129081 AAAAGATATTCATAGTGTCCAACAACATCTTAACTCAATACAAAGAATTA 129032
560
129031 CATCAGCTGTCCAAGTTACCTCAGATGTGACTTTAGTTGTGGAAGT 128982
560
128981 TACGGTGAAGTTGACTCCACACAATCCAATGCTGGTGCACACTCCTGTTA 128932
560
128931 CATAGATATTGACATTCATGAATACTTTCAATTACTCTTCTAAGATT 128882
560
128881 GTTACTGATAGACAGTTGGTAGATGCTGTTGACAACAAGATGTGCCAACT 128832
560
128831 TTGCATTCATATTTTAAAGATTGGGAACAAGATCTTCCTTCAAAATCA 128782
561Arg..AsnLeuLeuTy 565
128781 GCATTACAAAAAATATATGTAAACCATGTCAATTGCAGGAACCTTGCTCTA 128732
565 rIleTyProGlnSerLeuAsnPheAlaAsnArgGlnGlySerAlaArga 582
128731 TGTTTACCCAAAAAGCTTGAAGTTCAGTAGTGTGAGGCTCAGTGAGGA 128682
582 snIleThrValLysValGlnPheMetTyrgLysPProSerAsnAla 598
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

128681 ACATCGCTGTCAAAAGTCCAGTACATGCGCAGAGAGGAGACCCAGTCAAGCT 128632
599 Met.....
128631 TTGCAGGTGACGAGAAGTCAAGACTACCAAGTGATTAATAATACAATGTT 128582
599
128581 CTAAGAAAAACAGCACCGAATCTGTAATTCTGAGAATTTACCTGGCTTTAG 128532
599
128531 TAATTCTGCAGTAAGTAAAAAACAATCCCTAATGGAATTTGGTTATT 128482
600Pro..ValIlePheGlyLysSerSerCysSerGluPheSer 612
128481 TTTATTTTCACTAAGCTCATTTTGGGCAAGTCAAGTTGTGCTGAGTTCCACA 128432
613 LysGluAlaTyThrAlaValValTyThrHisAsn..... 623
128431 AATGAGGCATACAGTCCCATCATCTACACACAACAAGTAAGGCTTTCCAGA 128382
623
128381 CTTAACCTGTGTAACTCCTTTGTGTTTTCCTTAAATTCGTGTTTAA 128332
624Arg..SerProAspPheHisGluIuI 632
128331 CTCTTTTTTTGTAAAAAAAAGCTCCCTGAGTGTATGAGCAGA 128282
632 IeLysValLysLeuProAlaThrLeuThrAspHisHisLeuLeuPhe 648
128281 TGAAGATGAAGATTCTCTGCCAATCTGACTGACAATCACACACCTCCTGTT 128232
649 ThrPheTyrHisValSerCysGlnGlnLysGlnAsnThrProLeuGluTh 665
128231 ACCTTCTACCACATCAGCTGCCAGCCCAACAACACACTCCTCTAGAGAC 128182
665 rProValGlyTyT..... 669
128181 TCCTGTGGCTACACTGTGAGTTCACTCTAGTCAGTCTACATTTTATACT 128132
669
128131 CCCACCCACAGGCCACGTTTGATTAATCTCTGTGTGTCATGCCATCTTT 128082
670Thr..TrpIleProMetLeuGln 676
128081 AAAAGATCTGTGCTTTTCTTTCGACTTGTAGTGCATCCCTCTGATGCAG 128032
677 AsnGlyArgLeuLysThrGlyGlnPheCysLeuProValSerLeuGluLy 693
128031 CACGGCCGACTGAGAACAAGGCTCCTCAGTTTACCTGTCTCAGTGGAGAA 127982
693 sProProGlnAlaTyrSerValLeuSerProGluVal..... 705
127981 GCCCCACCGAGCTACTCTGTCTACACCCCTGATGTAAAGCAAGACTGGG 127932
705
127931 GATATCAGTTTGCTGAAATGACCCTTGCGCTTTTGACTGGCTGTAA 127882
706Pro..LeuProGlyMetLysTrpV 713
127881 TGAGATGGAATTTCGCTGTGATGTAGTTCAAGCTTCCAGGAGTGAAGTGG 127832
713 aLAspAsnHisLysGlyValPheAsnValGluValValAlaValSerSer 729
127831 TGGATATATCACAAGAGAGTGTTAGTGTGAAGTGACAGCAGCTTCTTCG 127782
730 ILeHisThr..Gln..... 733
127781 GTTCATACTCAGGTACACAGCGCTAAATAATTGTGCTTTCATCTATCAT 127732

733 733
127731 CTGACCAGCATTTCAAGCATCCATTTGATCTTAATGTTATAGTAA 127682
734 AspProTyrLeuAspLysPhePhea 742
127681 TATTAACTAGTCTGTAACTGTAGGACCCCTTCCTGGATAAGTTCTTCA 127632
742 lAlauValasnAlaleuAspGluHisLeuPheProValArgIleGlyAsp 758
:::||||| ::|||::|||:: ||||| |||:: |||
127631 CGCTCGTATGTCCTGGAGAGTACTCTCCCTTCGCTGAAGGAT 127582
759 MetArgIleMetGluAsnAsnLeuGluAsnGluLeuLysSerSerIleSe 775
::: ||| |||::|||::||| |||||::|||:::
127581 GTGATCATTAAGTGAACATGAGGACAGAGTTAAAGACCACCATGGT 127532
775 rAlaleuAsnSerSerGlnLeuGluProValValArgPheLeuHisLeuL 792
|||||:::~::~: |||:: ||| ||||| |||
127531 GGCACGTGAGGGGCCCTGTGGATACCTGTGTACAGCTTCCTCCACCAGC 127482
792 euleuAspLysLeuIleLeuLeuValIleArgProProValIleAlaGly 808
|||||::|||::||| |||::~::~: ||||| |||||::|||
127481 TGCTCAACACAGATCATTCAACTCATGCTTTTCCACCGGTCAATCGGAGGC 127432
809 GlnIle.Val..... 811
127431 CAGATTGGTAAATGAGCTCTGAAGACCGCTGCTGCTACGTCTACAA 127382
811 811
127381 TTGTGATTTGGAGTGTGCTTCTGTGAACAAGGCTGAAGGGCTCATTTCTA 127332
811 811
127331 ATAGTGGGTTAAATAGTTGATGTAATTGATCAAAATTACTGGATGGAAA 127282
811 811
127281 ATAATTGACATTTTGTCTCATACACATAAAACACTGACTTCATGTCAA 127232
811 811
127231 TGCTAAACCATGTTGTTTAAATAATCCTACAATCTTGATTTCCAACCTC 127182
812AsnLeuGlyGlnAlaSerPheGluAlaMetAlaSerIleIl 825
127181 CTGACAGTCAACCTTGGCAGAGACTGCTTTGAAACTATGGCTTTACTGGT 127132
825 eAsnArgLeuHisLysAsnLeuGluGlyAsnHisAspGlnHisGlyArgA 842
:||||: ||||| ||||| |||||::~::~: ||||| ||||| |||||
127131 GAACCAGACACAAAAAACCCTGGAGGGAACACGAGACCATCATGTGCGCA 127082
842 snSerLeuLeuAlaSerTyrIleHisTyrValPheArgLeuProAsnThr 858
||::||| |||||::||| ||||| ||||| |||||::~::~: ||||| ||||| |||||
127081 ACAACCTGCTGGCTTC.TACATCCAATACTGCTTTCGCCCTGCCCACTGCT 127033
859 TyrProAsnSerSerSerProGlyPro..... 867
||| |||::|||:::
127032 GAACCATC...AGTCCCCCAACAGGTGAGACCTCGCACCACTGTGCA 126987
868GlyGlyLeuGlyGlySer..... 873
126986 CAGCAGATTATTGGTTGGATTGCTGGCAGTGTCTGTGAGACCGGCCCGT 126937
873 873
126936 TGATCTGCCCTTATCTTCTCTGTGGCTCTCATCCACGTTTGAGAGAAACC 126887
874ValHisTyrAlaThrMetAlaArgSe 882
126886 AGTTCACAGTCCCATCAGATGCCCGGTGACGTACGCAACCATATCGAGGGC 126837

882 rAlaValArgProAlaSerLeuAsnLeuAsnArgSerArgSerLeuSera 899
::: |||||::||| |||||::||| |||||::||| |||||::||| |||||
126836 AACTGGACGCCCAAGCAGACCTGCACCTGTCCCGCTCCAAAGACATCAGCA 126787
899 snSerAsnProAspIleSerGlyThrProThrSerProAspAspGluVal 915
||||| |||||::~::~: ||||| ||||| ||||| ||||| ||||| |||||
126786 ACTCCAACCCAGACCTGGCCAGTTCACCAAGTTTCTCCGATGAAGAGGTG 126737
916 ArgSerIleIleGlySerLys.GlyLeu..... 924
::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
126736 CAAAGATCATAGGAAGCAAGGTGAGGCTCTTCTGGCTAATGAAGCTTTT 126687
924 924
126686 GTCTTTTATACAGCTTAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 126637
924 924
126636 GTGTGTGTGTGTTAAAGATTGAGAACATTAATCAGTAGAATCTGTTA 126587
924 924
126586 CATACTCTGAAGAGTGAGATTTCTCTCCTGTTTCAACAACAACACAGC 126537
924 924
126536 TTTTTCACAACAAGTAAGTTTTTGGGGGTAATTCTTGCCCTCCCAAAA 126487
924 924
126486 ACTGAATGTGACCTTCTGGGCTAGGCCCTTCCACACACAGTGTCTTAC 126437
924 924
126436 ATGTGGCTGGCCCTTTGTGTCAAGACTGTGTCTCTCTTGCGCAATATTCA 126387
924 924
126386 TCCACCAATTATGCCCATTCGCTCTTGTGATATATATATATATATATAT 126337
924 924
126336 TTGGGTTAAGTGTGAGAAAGAGCTGGAACAATCCCATATCAGGGAATG 126287
924 924
126286 ACATCATTTAGTTGACCCCGCCCGTGAGCTGATCTCATTTCTCGCTTCT 126237
924 924
126236 TTTAGAGTATGTGTTAATATGATTTTTTATGAAGTGCAATTATGCTT 126187
924 924
126186 TGTGTTTTTGAACATGTCTCTGAGCCCTTGATTCCTGCCCCAGGTGAG 126137
924 924
126136 ATTCTGTGTGTTTAAAGCTTATTTCTGTGCGGCCCGCCCTTTTCT 126087
924 924
126086 TTGTTTTTCCACCATTTTATCCTCCAATAATTAATTAATTAATTAATTA 126037
924 924
126036 CTTCCTTCACTACGTACCCCTTGCCCTGTCGCACTCTGTGTTTGTGT 125987
925AspArgSerAsnSerIlePVal 931
125986 TTGTCTTCGGTGGCTTATGCGAGGGCATTTGACCGCTCCCACTCCTGGGTT 125937
932 AsnThr.....GlyGlyPro..... 936

125936 AACCTGCTTATGCCCGGGGTTCCAGATCATGCTTCGCCGGAACCC 125887
937 LysAlaAlaProTrpGlySerAsnPro.... 945
125886 CAACCTCCAGCTGTGATCTGAAGCAGGTGCCAATCCCTCCTCCGTCCT 125837
945 945
125836 CCTACCTACCTGCCGTCGCCACGTTCTAAACAACACTCTTGAGTCATTGT 125787
945 945
125786 TTGCTGCATAAGCTTCCATCTGCTTTTCATGCTGCCACTCCTCTTATC 125737
945 945
125736 TGTGAATACTACTTATCTGTGTTCAACCATTCATCCTCTCAACCACCAATAA 125687
945 945
125686 CAGCCCAATCTGTTGTGAGTGCCTGCCTCGTGCAGAGTTTCTACTAGC 125637
945 945
125636 CTCACACTCATCGTACAGAGTCTGTGTCACCCACTAATTTCCATTCTAT 125587
946 SerPro.SerAlaGluSerThr 952
125586 CCATTTTCACTACATTTTGTCTGTGTGTTCAACCTCACCATTCACACCATGT 125537
953 GlnAlaMetAspArgSerCysAsnArgMetSerSerHisThrGluThrSe 969
125536 CAGACACAGCAGCCGACAGCTGCATTCGATGCTGCTTTCCTGGACAGCAT 125487
969 rSerPheLeuGlnThrLeuThr.GlyArgLeuProThrLys.Lys.... 983
125486 GGCCCTGTTTTCAGTTCACAGGACACATTTGCCAAGAAGGTAAAACTGT 125437
983 983
125436 CTCTACTATTAACAGAATACAGAGTGTCCAACCTCAACGAGATGCATTTT 125387
983 983
125386 GTGCTTCACATGATTCATTTTCATCTGTCATCATCATCATCGTACATCA 125337
983 983
125336 CCTCATATGTAGGTTAAATGCTGACACTGATACTAAACATTTTAAACCA 125287
983 983
125286 AGCTATGTATTTATTTCAATTATTCATTTTAAATAATATACTTAAT 125237
983 983
125236 AACACTTAATCCTCACACAACGCTGATGTGTTTAGAGTTTCTGTTTTC 125187
983 983
125186 AATTTAAATCTAATTCATATTTTAAACATATATATATATGACAAAATTC 125137
983 983
125136 TAAGTTAAGCAGTGATGTGAAAACTTCTAGTGTGTTGTTCTCTGTA 125087
983 983
125086 AATGTATATTCACGTTTATGTGATGGAATATTTAGATTTTGTAAATAT 125037
983 983

125036 ATGACTGTTCTGTATTTTGAGCCGCAAAATCACTTTCAGAGTTGTCAAAAC 124987
983 983
124986 ACCTTTGTGAATGCTCTACAGGTAAATGGGTAAATTAGTAACCTGTGTA 124937
983 983
124936 TCTAAGGAATCTGTAAACATTACAGTAAGAAGTTGGCTAAATTTAT 124887
983 983
124886 GGCTTCTAGCATCATAGTAATAACAGCAAAACATTATAGTCAGCAAG 124837
983 983
124836 CGATAATTAGTTAGGCACCTGCTGATGTTTCTCAACGTCGCATATT 124787
983 983
124786 GCACCAAAGTATGATGATATATAACTGACTTTATTGTTGCTTCGTGTG 124737
984 ..LeuPheHisGluGluLeuAlaLeuGlnTrpValCysSerglyse 999
124736 CAGTTGCTTCATGAGGAGTGTGCTTTCAGTGGGTGTCAGCACCAACCCC 124687
999 rValArgGluSerAlaLeuGlnGlnAlaTrpPhePheGluLeuMetv 1016
124686 AGTGAGGAAGCAGCGCTTCAGCACGCCGTGTTTCTTCCAGCTGATAG 124637
1016 al..... 1016
124636 TGAGCATGCTGTGTGTGTCGTCAAAATCTCCATAAACACATTTTAAAA 124587
124586 TAATAGTCTTGAAGTCCCATCTCTTCTTCCCTTCTTTTCTTACTCTTTT 124537
1017Lys..SerMetValHisHisLeuTyrPheAsnAs 1027
124536 TGACTCTTCCCCAGACACAAGAGATGTCATCATCACTTATTTTCGACCTC 124487
1027 pLysLeuGlnAlaProArgLysSerArgPheProGluArgPheMetAspa 1044
124486 AAACGTGATCTACCGAGACGCCAACGATTTCCACAGACCGCTTGTGATG 124437
1044 sPLeuAlaLeuValSerThrIleAlaSerAspIleValSerArgPhe 1060
124436 ACATCGCTGCGCTTGTTCGTCATCAGTGCAGATATTTGCTGTCGATAC 124387
1061 GlnLys..... 1062
124386 CACAAGGTACTTGAGATTGCGGTACATTTCTTCAAAAGCTGAGGTTTTTT 124337
1062 1062
124336 TAGTGATGAATAAGCATTAAGTGCAAAACAATAATTGATCCGTTTAATAT 124287
1062 1062
124286 GTTGAACATCTCCTAAACATCCTCTTGCCCTAATTTTCACATGAACCTCCA 124237
1063AspThrGluMetValGluArgLeuAsnTh 1072
124236 ACTGTTTCTCCCTGTACAGAGATGTTGAACCTTGTGAGAGGTTAAATAG 124187
1072 rSerLeuAlaPhePheLeuAsnAspLeuSerValMetAspArgGlyP 1089
124186 CAGTCTGGCCTTCTTCTGAATGACCTGCTTCTCTCATGTGACCGAGGCT 124137
1089 heValPheSerLeuIleLysSerCysTyrLysGlnVal..Ser..... 1102
124136 TTGTTTCAACCTCATCCGCTCGTACTACAAACAGGTGCTGTGCTCTTCT 124087

1102 1102
124086 CTCCCTGTTGCTTATATCAACACTTTCCTTCACATATATTTGT 124037
1103 SerLysLeuTyrSerLeuPr 1109
124036 CTCACGTGCTGTGCGTTTAACAGATTGCCACAAGCTTCACACGGCACA 123987
1109 oAsnProSerValLeuValSerLeuArgLeuAspPheLeuArgIleIec 1126
123986 GAATCCAGCTCCCTGAATGCCCTGAGATGGACTTTATCCGTATCGTGT 123937
1126 ysSerHisGlnHisTyrValThrLeuAsnLeuProCysSerLeuLeuThr 1142
123936 GTAGCCATGAGCACTACGTATCCTCAACCTGCCGTGTTCACACTCCAGC 123887
1143 ProProAlaSerProSerProSerValSerSerAlaThrSer.Gln.... 1157
123886 CCTCCAGCCTCGCCCTCACCTTCACACCTCCACACAGTCACAGTGTGCT 123837
1157 1157
123836 ATTACATGCTGGAGGCTTAGTACAATTCATTACATTGCTCACATATCC 123787
1158 Se 1158
123786 AATATCTTTTCTTCTTCATTTCACCTTCTTTATTTTACGTTCCTAGAG 123737
1158 rSerGlyPheSerThrAsnValGlnAspGlnLysIleAlaAsnMetPheg 1175
123736 TTCAGCATTTTTCAGCATGTTGCAAGATCAAGAGTGGCCACAATGTTCG 123687
1175 IuLeuSerValProPheArgGlnGlnHisTyrLeuAlaGlyLeuValLeu 1191
123686 AGCTCTCCGTCCTGTTCCGCATCAACATTTCTCTCTGCGCTGTGTTA 123637
1192 ThrGlnLeuAlaValIleLeuAspProAspAlaGlu.Gly..... 1204
123636 TCAGAGCTCTCTCATCTCCTGATCGTGAAGGTAGGAGTGAA 123587
1205 LeuPheGly..... 1207
123586 TCTTACTTCTTTTTTTATGTCGTGGCTAATTGGACCATGTCTAATTGAT 123537
1207 1207
123536 ACTGAAATGCTGTTATGCTTCAAGTCCAGTACATGAAAACTGAGTTGTT 123487
1208 LeuHisLysLysValIleAsnM 1215
123486 GGTTTTTCCTACACACAGCATTTCTTCTCCATAAAAAGCCATTAGTG 123437
1215 etValHisAsnLeuLeuSerSerHisAspSerAspProArgTyrSerAsp 1231
123436 CTGTACACTCCCTGCTGTGCAGCCATGATGCAGATTCTCGCTACAAGGAC 123387
1232 ProGlnIleLysAlaArgValAlaMetLeuTyrLeuProLeuIleGlyI 1248
123386 CCTCAGGTCAGGCGCCATGTGTGCTCAGCTTTTACCTGCCCTCATTTCCAAT 123337
1248 eIleMetGlnThrValProGlnLeuTyrAsp..... 1258
123336 TGTCAATGAGACACTGAATCAGCTTCATGACTTCTGTGTGTCAGTGTACAC 123287
1259 PheThrGlnThrHisAsnGln 1265
123386 CCCCACACATTTTCATAACTTCGCAAAATGTTCACTTCGTGGCAATTTTAAG 123237
1266 Arg..... 1266
123236 ATGTTAATTTCATGCTGTGTGTATGTTTTGTGTAGACTCTCCCTG 123187

1267 .GlyArgProIleCysIleAlaThrAspAspTyrGluSerGluSerGly 1282
123186 CTGGGTCCGGATGCCCTCAGCCCATGACGATGCTGACCTGGATAGTGCC 123137
1283 SerMetIleSerGlnThrValAlaMetAlaIleAlaGlyThrSerValPr 1299
123136 AACACCATCAGTCAGTGTGTGCCATGGAATGCGCGCTCCCTTTTACC 123087
1299 O..... 1299
123086 ACATGCCAAAGCCAACTCCTTGGCGCTACCCACAGCGGTGATGATAATG 123037
1300 GlnLeuThrArgProGlySerPhe..... 1307
123036 GCTGGGGTCAAGTTGTTC.CCTGGGCATTACAGAAACAGGTCACAGATC 122988
1307 1307
122987 CCACTGTGTGGAGACTGTTGCCCTCAGCTTGACAGAACACACACACTGG 122938
1307 1307
122937 CCTTCTGTAACTCCACTAAAAAACAACAATAATGTGCATTCTTTCAC 122888
1308 LeuLeuThrSerThrSerGly 1314
122887 ATTGCTCATCTTCCCTCAAGGATAACTGCTGTGTGTGCCATAGGTGGC 122838
1314 1314
122837 TCTCTGCCCTCTATAAAGTGTCTTTCATGTTTGACATTGAAGGGCA 122788
1315 ArgGlnHisThrThp 1320
122787 ACTTAGCGGTGTGTGTGTCTTTCGCCAGCGCTGCCGCCAGTCCAGCTCTC 122738
1320 heserAlaGluSerSerArgSerLeuLeuIleCysLeuLeuTrpValLeu 1336
122737 TGTCAAGAGAGTGTAGCGGAGCTCTGCTGTGCTTCCCTGTGGGTGCTG 122688
1337 LysAsnAlaAspGlnThrValLeuGlnLysTrpPheThrAspLeuSerVa 1353
122687 AAGAATGCAGATGCAGCTCTCTGAGCGCTGGGTGTCTGATTTGTACAT 122638
1353 IleGlnLeuAsnArgLeuLeuAspLeuLeuTyrLeuCysValSerCysP 1370
122637 GCTGCAATATCAACCGCTGCTGATCTGCTGCAGCTCTGCATCTCTGTT 122588
1370 heGluTyrLys..... 1373
122587 TTGAATATAAAGTAAGTACTTGTGTGATTTGTTCCTTTTGTGTTTACGTGC 122538
1374 GlyLysLysVa 1377
122537 TTGAGTCTGTGACATGTTTACTTTTGTATTTAATAGGGGAAAAAGGC 122488
1377 IPhGluArgMetAsnSerLeuThrPheLysLysSerLysAspMetArgA 1394
122487 TCTGAGAGAGATTAAATAGCCTGACATTTAAAAAATCCCTGACATGAAGG 122438
1394 IalysLeuGlnGlnAlaIleLeuGlySerIleGlyAlaArgGlnGlnMet 1410
122437 CACGCTTGAGAGAGGCCATACTGGGAACCATTTGGGCTCGTCAGAGATG 122388
1411 ValArgArgSerArgGly..... 1416
122387 GTTTGTGCTTGACAGAGTAATAGTGTCTGTCTGTGTGTGTGTGTGTG 122338
1416 1416
122337 TGA 122288
1417 GlnLeuGluArgSerProSerGlySerAlaPh 1427

```
122287 TTCAATGTTAATCTCAAAATCTTTTAGAGAGAGTCCC.....TA 122250
::: ||||| |||||
1427 eGlySerGlnGluAsnLeuArgTrpArgLysAspMetThrHisTrpArg 1444
::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
122249 TAGCAGTCAAGAGAATGTGAGGTGAGAAAGACGTCACTGCGGCC 122200
1444 lAsnThrGluLysLeuAspLys..... 1451
||||| ||||| |||||
122199 AAAATACAGACAGAGTCGACAAATATGTTTTTTTAACTTTATTG 122150
1451 ..... 1451
122149 CATTAATGCCAGTTACTGTAAAGATGTTCCATGCGGTATAGCAACTG 122100
1451 ..... 1451
122099 AGTGTTCATCACGGGTGGGATACATCTGCTTTAATTTCCAAG 122050
1452 .....Ser.A 1453
::: :
122049 TGCTTGAATGCTGATGTGAACCTAAGCAATTAATCTTTATAGGACTA 122000
1453 rGAlaGluIleGluHisGluAlaLeuIleAspGlyAsnLeuAlaThrGlu 1469
::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
121999 AAGCTGAGATGAGCAGCAGAGAGTCGTTGGATGGAACCTTATCAACCGAG 121950
1470 AlaAsnLeuIleIleLeuAspThrLeuGluIleValVal.Gln..... 1483
|||::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
121949 GCTTCTTAATCGTCTCTTGACACCGTGAAGATTGTAGTCAAGGTAGACAA 121900
1483 ..... 1483
121899 TGCTTTCAAGCTGTGAAGTGCAGAGATTATTGAGAATTAAATTCTGTT 121850
1484 .....ThrValSerValThrGluSerLysG 1492
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
121849 TCTTTCGCTTTGATTTTTCAGACTGTGGTGGCATCAGAGCTGAAGGA 121800
1492 uSerIleLeuGlyGlyValLeuLysValLeuLeuHisSerMetAlaCysA 1509
|||::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
121799 AAGTGTCTGGGTGGAGTAATGAGAGTCTCTCCACACAGCATGGCAGGAA 121750
1509 sngInSerAlaValTyrLeuGlnHisCysPheAlaThrGlnArgAlaLeu 1525
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
121749 ACCAGAGTCCCTCTCTCTGACACTGCTTCACTACACAAAGAGCTCTA 121700
1526 Val..... 1526
|||
121699 GTTTTCAAGTACATAAATTATTTTGTTAATTGTGATGATATTTT 121650
1526 ..... 1526
121649 CTCTTATCTTTTACATGTTGCAGCATATTACATCATAGATTAGATGC 121600
1527 .....Ser.Lys.... 1528
||| |||
121599 ATACATACATATAGATAGACATGTTGTGACCACTTGAAGTGGAAAGTGA 121550
1529 .....PhePro 1530
|||||
121549 AGGCAACTGAATACATGTGTATTATATACCTGCTGCTGTAGTTCCA 121500
1531 GluLeuLeuPheGluGluThrGluGlnCysAlaAspLeuCysLeuAr 1547
|||::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
121499 GAGATGCTGTTGAAGAGGATACAGAACTTTGTGCCGACCTATGTCTGCG 121450
1547 gLeuLeuArgHisCysSerSerSerIleGlyThrIleArgSerHisPros 1564
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
121449 TCTTCTGCGTCACTGCAGTAGCAGCGCTCACTCTGTCAAGAGTCAAGCCT 121400
1564 erAlaSerLeuTyrLeuLeuMetArgGlnAsnPheGluIleGlyAsn... 1579
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
121399 CTGCCCTCTCTTTATCTGCTCATGAGACAGAACTTTGAGATTGAAATGTA 121350
1579 ..... 1579
121349 AGTAATATACTTTCATGTTCATGATGTAATATATACTTTCATTATGTTCA 121300
1579 ..... 1579
121299 ATACCTACTTTAGGAAGTTTACTTGAGATTTACTTAAAGAGTGAGACAAA 121250
1580 .....AsnPheAlaArg 1583
||||| ||||| |||||
121249 GGAAGAAATTTACATCCCTTTTATTTCCTTTTGACAGAACTTGTCTCGG 121200
1584 ValLysMetGlnValPrometSerLeuSerSerLeuValGlyThrSerG 1600
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
121199 GTAAGATGCAGGTCAACCATGTCTACTGTCACTGCTGTTGGGACATCAAG 121150
1600 nasnPheAsnGluGluPheLeuArgSerLeuLysThrIleLeuThrT 1617
::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
121149 AAACCTCAATGAGGAACATCTTGACACTCGTTGAAAACCTATCTTGACAT 121100
1617 yRAlaGluGluAspLeuGluLeuArgGluThrThrPheProaspGln.V 1633
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
121099 ATGCTGAAGATGATGTGGAACACTACGTGACACACCTTCCCGCAGAGGTA 121050
1633 al..... 1633
:::
121049 TTAATGGCCAGTGTAGAAGATCCCATTTGTAAGATGCATGAATGGTTG 121000
1634 .....GlnAspLeuValPh 1638
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
120999 AGATTTTATTCATTCCTGTGTGTAACACTCAGGTCAGGATTTGGTTT 120950
1638 eAsnLeuHisMetIleLeuSerAspThrValLysMetLysGlnHisGlnG 1655
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
120949 CAATTTCACATGATTTCTCACCGATACGTGTGAATAATGAAGAACACCAAC 120900
1655 lAspProGluMetLeuIleAspLeuMetTyrArg.Ile..... 1667
::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
120899 AGGATCCAGAGATGCTTATTGACCTGATGTACAGGTAATGCGATGCACAT 120850
1667 ..... 1667
120849 TAATAATCAAAATATAGTATCAGACTTAATATCTGCTTTTGTGTTAAATC 120800
1668 .....AlaLysGlyTyrGlnThrSerProGlu...ArgLeuThrTrpLe 1681
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
120799 AGGATTCGAAGGGCATATCAAAACTCTCCGACCTGCGCTGACTTGCT 120750
1681 uGlnAsnMetAlaGlyLysHisSerGluArgSerAsnHisAlaGluAla 1698
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
120749 CCAGAACATGGCAAGAAACACTCTGAGAGGAGAAACCATGCTGAAGCTG 120700
1698 lAglnCysLeuValHisSerAlaAlaLeuValAlaGluTyrLeuSerMet 1714
|||::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
120699 CCCACTGCCCTGCTCCACAGTGCAGCACTGGTTGCAGAAATACCTCAACATG 120650
1715 LeuGluAspArgLysTyrLeuProValGlyCysValThr..... 1727
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
120649 CTGGAGGATTGTGCTACTTGCCAATCGTTGCGTCAACCTTTCAGTGGG 120600
1727 ..... 1727
120599 TGAGAAAGCAGCTGACTTTAAGTATGTGATGCAGCAGCTAGTTGGCTGTCG 120550
1728 .....PheGlnAsnIleSerSerAsnVal 1735
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
120549 TTATTAATATGCATCATACACATACGCTTTCAGAACATTTTCACCAACGTC 120500
1736 LeuGluGluSerAlaValSerAspValValSerProaspGluGlu 1752
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
120499 CTGGAGGAGTCTGCTGTCTGTATGACGTGTTGTCTCCGGAAGAGAGGG 120450
```

1752 yIleCysserGlyLysTyrPheThrGluSerGlyLeuValGlyLeuLeuG 1769
|||||:::|||||
120449 GATCTGTGCTGGAAGACTTTCAGCGAGTCTGGTCTAGTGGCCTCTTG 120400
1769 luGlAlaAlaAlaSerPheSerMet..... 1777
|||||
120399 AGCAAGCAGCTGCTCTTCAACATGGTACTGCTCACATTTGTCACACC 120350
1777 1777
120349 ATGGTGACAGAGTGGCTGTGTCTCCCATGGCTGTGGGTTTATATGAT 120300
1778Ala. 1778
120299 TGGCGCTTGTGTAAACTGGCCTACCAATGTGTGTTGACTGCAGGCT 120250
1779 GlyMetTyrGluAlaValAsnGluValTyrLysValLeuIleProIleH 1795
:::|||||
120249 GCCATGTATGAAGCCATTAATGAGGTTTACAGATTTCTGCTGCCCATCCA 120200
1795 sGluAlaAsnArgAspAlaLysLysLeuSerThrIleHisGlyLysLeuG 1812
|||||
120199 TGAAGCCAACAGACAGACTTAAGAAGTTGGCTACAGTCCATGGGAAGCTGC 120150
1812 lngluAlaPheSerLysIleValHis..... 1820
|||||
120149 AGGAAGCTTTTAACAAGTCTACAACCAAGTGAAGTTGACGGCTGAATGAA 120100
1820 1820
120099 CGACAGATATCATATTACTCTGAACCTTCTCCTAATGTTGATTTTCTTT 120050
1821GlnSerThrGlyTyr.Glu..... 1826
|||||
120049 TGTTTTCACCTTTCAGAGTTTCAGAGTGGAGTAATGCTTATTTTAAAAA 120000
1826 1826
119999 AATGACTGATGAGCACTTACTTTATTAAGTCACATTAAACATTTCTTGC 119950
1826 1826
119949 TTCTTCGCTTATCTCCCCCTTAATCCCTTTGAATGACAATATGGGT 119900
1826 1826
119899 GAGTTTAAACTTATTTTATCTTAATAAATATGAGTTTGTCTTTAAA 119850
1827Ar 1827
119849 GCTGTGATTTAAATATCATCATCTTACCATCCCATCTTCAGTGGTTAGAG 119800
1827 gmetPheGlyThrTyrPheArgValGlyPheTyrGlyThrLysPheGlyA 1844
|||||
119799 AATGTCGGGACCTATTTCGGGGTGGTTCCTTGGTTCGCTTTGGAG 119750
1844 spleuAspGluGlnGluPheValTyrLysGluProAlaIleThrLysLeu 1860
|||||
119749 ACCTGGACGACAAGAGTTTGTCTCAAGAGAGCCATCCATCACTAAACTG 119700
1861 AlaGluIleSerHisArgLeuGluGly..... 1869
|||||
119699 GCAGAGATCTCTCACAGACTTGAGGCAATGTTGACTCACCGCAGCCTCAAG 119650
1869 1869
119649 AAAAGCAGTAATTACCGCTTCCCAAATATATAGGCTTTCATATACGTGT 119600
1870PheTyrGlyGluArgPheGlyGluAspValVal 1880
|||||
119599 GTTTTGTCTGTGCAGGAGTTTACTTCAGAGCAGTTCCGAGATGACGTGGTT 119550

1881 GluValIleLysAspSerAsnProValAspLysCysLysLeuAspProAs 1897
:::|||||
119549 AAAATTTCAAGGACTCCAAATCCAGTAGACAAAGAACAACTGCATCCCAA 119500
1897 nLys..... 1898
119499 CAAGGTGCCGTGATATTAGGCTCATGCTGAATGCTCTGCTCAGTTT 119450
1899AlaTyrIleGlnIle 1903
119449 GTGTAATTATTCTTATAGTGTATGTACTCTTCAGGCTTACCTCCAGATC 119400
1904 ThrTyrValGluProTyrPheAspThrTyrGluMetLysAspArgIleTh 1920
|||||
119399 ACCTACGTCAAGACCATTCCTCGACACACGTATGAGCTAAAGAAAGAGTCAC 119350
1920 rTyrPheAspLysAsnTyrAsnLeuArgArgPheMetTyrCysThrProp 1937
|||||
119349 CTACTTTGACAAGAAGACTACAACTGCGTAACCTTCATGTACTCTCCCT 119300
1937 heThrLeuAspGlyArgAlaHisGlyGluLeuHisGluGlnPheLysArg 1953
|||||
119299 TCACCTTGACGGGGCGAGCAGCAGCGCAACTGAACGAGCAGTACAAACGT 119250
1954 LysThrIleLeuThrThrSerHisAlaPheProTyrIleLysThrArgVa 1970
|||||
119249 AAAACCATCTCACAACGCTCTCACGCTTCCCTTACATCAAGACACGAGAT 119200
1970 lAsnValThrHisLysGluGlu..... 1977
|||||
119199 CAATGTTATTCACAAGGAGGAGGTGTGTTCTGAGAGAGCGCAATTCTCT 119150
1977 1977
119149 AGGTGATGGGATGCGCTCCAGCAGGGGTATAAATGTGTGCTGTGTTCTC 119100
1978 ...IleIleLeuThrProIleGluValAlaIleGluAspMetGlnLysLy 1993
|||||
119099 CAGATCATTTCTCGTTCCCATAGAGGTGGCCATGAGGACATGCAGAAAGAA 119050
1993 sThrGlnGluLeuAlaPheAlaThrHisGlnAspProAlaAspProLysM 2010
|||||
119049 GACACAAGAGTTGGCCTTGGCCACAACAACGAGCCTGCAGACTCCACAGA 119000
2010 etLeuGlnMetValLeuGlnGlySerValGlyThrThrValAsnGln... 2025
|||||
118999 TGTTCAGATGGTGTCTGCAGGGCTGCGTGGGCACCACTGTCAACCAGGTG 118950
2025 2025
118949 CCTGTTTACACACGTGTGTGATGTATATCCAAATGAAGTGTCAAAATGC 118900
2025 2025
118899 TATCCCTTTATTAATGGGCTTTAAAGTCTGTGCCAGTTGTTTATGTAG 118850
2026Gly 2026
118849 CGCCACACATTCCTCGAGAAGATAAATGACGCTCTTCCCTTCACAGGG 118800
2027 .ProLeuGluValAlaGlnValPheLeuSerGluIleProSerAspProL 2043
|||||
118799 CCCCCTTGAGGTAGGCGAGGTCTTCTTCCGACATTCCTGATGATCCAA 118750
2043 ysLeuPheArgHisHisAsnLysLeuArgLeuCysPheLysAspPheThr 2059
|||||
118749 AGCTGTTTCGCCATCACACAAGCTGCGCCTGTGCTTTAAAGACTTTCAC 118700
2060 .Lys..... 2060
118699 AAGAGGTGATGACACCGGTGTGACTCAGATGCAGACTGCTGTGTAATAAT 118650
2061ArgCysGlu 2063


```

118649 AAAACAAACACACACTGCTGTCTCTGCATTCCTCACAGGTGTAG 118600
2064 AspalaleuArgLysAsnLysSerLeuIleGlyProValGlnLysGluTy 2080
|||||
118599 GATGCCCTGAGGACAGATAAAGCCCTCATTTGGCCGATCAGAAGAGTA 118550
2080 rGlnArgGluLeu 2084
|:::|||||
118549 CCACAGAGAGCTG 118537

seq_name: gb_htg:AC017375

seq_documentation_block:
LOCUS AC017375 73515 bp DNA linear HTG 09-DEC-1999
DEFINITION Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***, in ordered
pieces.
ACCESSION AC017375
VERSION AC017375.1 GI:6553611
KEYWORDS HTG; HTGS_PHASE2.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 73515)
AUTHORS Adams,M. and Venter,J.C.
TITLE Direct Submission
JOURNAL Submitted (09-DEC-1999) Celera Genomics, 45 West Gude Drive,
Rockville, MD, USA

COMMENT This sequence was identified as CDM:10210477 by the submitter.
For more information on this record e-mail to fly@celera.com.
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

FEATURES
source 1..73515
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
BASE COUNT 21627 a 14907 c 14713 g 22268 t
ORIGIN

alignment_scores:
Quality: 4700.00 Length: 2339
Ratio: 2.892 Gaps: 57
Percent Similarity: 69.474 Percent Identity: 44.164

alignment_block:
US-09-737-246-2 x AC017375/rev ..

Align seg 1/1 to reverse of: AC017375 from: 1 to: 73515

4 ArgArgAlaPheAlaGlnLysIleSer..... 12
:::|||||:::|||||:::|||||
43211 CAGCGCTCGTTGTGCAAAAAGTTCCAAAGTAAGTTTTCACATAAGGAA 43162
12 ..... 12

43161 ATAGTTATTTTAAACAAGAAAGCGCCCTTAATGCTTGATACATATGTA 43112
13 .....ArgThrValAlaAlaGluValArgLys 21
||| |||||:::|||||
43111 CATATATCAATTTGTAAATCCCTAGGACAGTCGGCTGCGGATGTACGAAAA 43062

22 GlnIleSerGlyGlnTyrrSerGlySerProGlnLeuLeuLysAsnLeuAs 38
:::|||||::: |||::: |||:::
43061 AATGTTCTGTGTTGCCATCTATCAAAAGCAATG.....GACCCATC 43021

38 nIleValGlyAsnIleSer.HisHisThrThrValProLeuThr..... 52
::: |||||::: |||::: |||
43020 ATGTGTGGTCTTCGCGTAAGACACATATTTTAAATATACAAATTTT 42971

```

53GluAlaValAspProValas	59
42970	CTGTTAAGGATTCCATTTTTCATAGATCTCCCGACAGAGCCCTTGA	42921
59	pleuGluAspTyrLeuIleThrHisProLeuAlaValAspSergIleProL	76
42920	TTACGAGAAGATTTCATCTCCACCATATGAACATCATCAACCGGATCCCT	42871
76	euaArgAspLeuIleGluPheProProAspAspIleGluValValTyrSer	92
42870	TGAACATATATTTGACTTTCCTCAAGGGATGTAAATCATTT	42821
93	ProArgAspCysArgThr.LeuValSerAlaValProGlu.....	105
42820	CCCCGAAAAATCAGAACTGTTGATCACGTTGTTCCCATGAGAAATTGTA	42771
106Glu	106
42770	TGTATGCACCTAAGTTATACACAGCTTACATAGCTACAGTCTCTTTAGG	42721
107	SergIleMetAspProHisValArgAspCysIleArgSergTyrThrGluAs	123
42720	TCGTGATTTGCCCATCAGCTACGTCACAAGATGTGTTAACTGCTATACGCCCC	42671
123	pTrrAlaIleVal.....I	128
42670	ATGGAAGGTGGTTGAATATAGCACAGCGCACCTTTCAGTCTTGTGCTATA	42621
128	IleArgLysTyrHisLysLysGlyThrGlyPheAsnProAsnThrLeuAsp	144
42620	TCCGCGAGCGTATCGACCGCGGAACC..ATAAGTCCCTGTGCGTACCAG	42574
145	LysGlnLysGluArgGlnLysGlyLeuProLysGlnValPheGluSerAs	161
42573	CAGGAATTCGAGATCGACAAGAT.....TTTTCGTCAATT	42539
161	pGluAlaProAspGlyAsnSergTyrGlnAspAspGln..AspAspLeuL	177
42538	TGAA.....GAGCGTTCGCTTTAAAGCGAGAGAGCTGCATCCCAAGTT	42495
177	ysArgArgSermet.....SerIleAspAsp...	185
42494	CTCGACAAATCCATTGCAAGTTTGGCTTCAGTAAGCTCTTGTACGACACT	42445
186	..ThrProArgGlySergTrrAlaCysSerIlePheAspLeuLysAsnSe	201
42444	TTGACCCCGCGTGGTCTTGGGGCTAGC....TTTGATCTTCGACGCTTC	42401
201	rLeuProAspAlaLeuLeuProAsnLeuLeuAspArgThrProAsnGluG	218
42400	GGTTAATGACCCACTAATATACCCAAATTGCTTGACAATGTACTCCGGAGC	42351
218	IuIleAspArgGlnAsnAspAspGlnArgLysSerAsnArgHisLysGlu	234
42350	ATATATGATCATCCAAATGAAGCACGCCGCGACGAAGATCGACAAATGGCT	42301
235	LeuPheAlaLeuHisProSerProAspGluGluGluProIleGluArgLe	251
42300	CTTTTCTCACTGATCCGGAATCTGAAGCGGAGGAAAAACATTTGAACGCCG	42251
251	userValProAspIleProLysGlnHisPheGlyGlnArgLeuLeuValL	268
42250	CCTGCTCGCAGAGATTCACAGTTGAACATATGGGTCAATCGCATTCAGGTTA	42201
268	ysCysLeuSergLeuLysPheGluIleGluIleGluProIlePheAlaSer	284
42200	ACTGCCTACAGCTTCGGCTTGAGCTGGAAGTGAAGCCAAATTTTGCATCG	42151
285	LeuAlaLeuTyrAspValLysGluLysLysLysIleSergLysAsnPheTy	301
42150	ATGGCTATTATATGATGCTAAGGAACCGCAAAAAATTTCTGAATAATTTCTA	42101

1449 uasplySerArgAlaGluIleGluHisGluAlaLeuIleAspGlyAsnL 1466
: |||||::: |||::||| :::: |||:::|
38736 GGGAAAAAGTGAGCATGAACTTGAAATTAGATTCATTCGAAAGGCTCT 38687
1466 euAlaThrGluAlaAsnLeuIleIleuAspThrLeuGluIleVal 1482
|||||::: |||:::| |||||:::|
38686 TGGCAACTGAAGTGCACCTTGACTTCTCGATACCTTGAATAATGTT 38637
1483 GluThrValSerValThrGluSerLysGluSerIleLeuGlyGlyValLe 1499
::: :::::|:|:|:|:|:|:|:|:|
38636 CAT.....GCTGCTGCCAACCTTTATCACAACTCTCTTGAACCTGTGCT 38593
1499 uLysValLeuLeuHisSerMetAlaCysAsnGlnSerAlaValTyrLeuG 1516
|||||:::| |||||:::|:|:|:|:|:|:|:|:|
38592 AAAGTGCTCTTCATCTTGTGTCGCCGAATCAATCTGTTGGCCCTGC 38543
1516 lnhisCysPheAlaThrGlnArgAlaLeuValSerLysPheProGluLeu 1532
|||::: |||||:::| |||||:::|:|:|:|:|:|:|
38542 AAAACTGTGTGCCCTCCCAACGGCGTTAATCTTAAATTCCAAACTTA 38493
1533 LeuPheGluGluGluThrGluGluCysAlaAspLeuCysLeuArgLeuLe 1549
|||||:::| |||||:::|:|:|:|:|:|:| |||||
38492 TTATTTGATGACGAGACCGATATATGTGCGGATTTGTCCTAATTTTACT 38443
1549 uArgHisCysSerSerSerIleGlyThrIleArgSerHisProSerAla 1566
|:::| |||||:::|:|:|:|:|:|:|:|:|:|:|:|
38442 GAAACATTTGGGTCCCTATTGTCAGGAATPACGATCACAGCAGCTGCAT 38393
1566 erLeuTyrLeuLeuMetArgGlnAsnPheGluIleGlyAsn..... 1579
38392 CACTATATTACTAATGAGACAAATTTGAAATTTGAAATGTAAGTTT 38343
1579 1579
38342 ATTTTATGTTAGCTTATTAACTAAATTAATGCTACTTATTCGTAT 38293
1580 .Asn.PheAlaArgValLysMetGlnValPrometSerLeuSerSerLeu 1595
::: |||||:::| |||||:::|:|:|:|:|:|:|
38292 AGAAGTTGCCCGAGTTAAGATGCAAGTGACGATGCTTTAAGCTCCCTA 38243
1596 ValGlyThrSerGlnAsnPheAsnGluPheLeuArgArgSerLeuLy 1612
|||||:::| |||||:::|:|:|:|:|:|:| |||||
38242 GTTGAACAAGTTCGGTCTTTAGTGAGCAATCCTTACGCCGTCGCTTAA 38193
1612 sThrIleLeuThrTyrAlaGluGluAspLeuGluLeuArgGluThrP 1629
|||||:::| |||||:::|:|:|:|:|:|:|:|:|:|:|:|
38192 AACCGTCTGTTTATGCTGAATCCGACTCCGACCTGCAGGATACTCTT 38143
1629 heProAspGlnValGlnAspLeuValPheAsnLeuHisMetIleuSer 1645
|||||:::| |||||:::|:|:|:|:|:|:| |||||
38142 TTCCTGAACAAGTACAAAGATTTGCTTTTAACTGCATATGATCCTGTCA 38093
1646 AspThrValLysMetLysGluHisGlnLysAspProGluMetLeuIleAs 1662
|||||:::| |||||:::|:|:|:|:|:|:| |||||
38092 GATACGTGTTAAATGAAGAGTATCAGAGAAGACCCAGAATGTGCTTGA 38043
1662 pleuMetTyrArgIleAlaLysGlyTyrGlnThrSerProGlu...ArgL 1678
|||||:::| |||||:::|:|:|:|:|:|:|:|:|:|:|:|
38042 CCTCATGAATCGTATGCCCAGGATACCAAAATAACCTGATCTACGAC 37993
1678 euThrTyrPleuGlnAsnMetAlaGlyLysHisSerGluArgSerAsnHis 1694
|||||:::| |||||:::|:|:|:|:|:|:| |||||
37992 TGACTTGGTTGAAAAATATGGCTTAAAAAACACCGCGAGCGAGCAAAATCAC 37943
1695 AlaGluAlaAlaGlnCysLeuValHisSerAlaAlaLeuValAlaGluTyr 1711
:::| |||||:::|:|:|:|:|:|:|:|:|:|:|:|
37942 ACGGAAGCAGCCATGTGTATGTACATGCTGCTTCTTAGTTCTGAATA 37893
1711 rLeuSerMetLeuGluAspArgLysTyrLeuProValGlyCysValThrP 1728
|||||:::| |||||:::|:|:|:|:|:|:| |||||
37892 TCTTAGCATGTGGAGTACACAAAAACATTTGCCCTGTGGAGCTGTAAGTT 37843
1728 heGlnAsnIleSerSerAsnValLeuGluLeuSerAlaValSerAspAsp 1744

|||||:::| |||||:::|:|:|:|:|:|:| |||||
37842 TTCACGAATTTCTCCCAACACACTAATGAGTGGCGGTATCGGATGAT 37793
1745 ValValSerProAspGluGluGlyIleCysSerGlyLysTyrPheThrG 1761
|||:::| |||||:::|:|:|:|:|:|:| |||:::|
37792 GTGCTAAGTCCCGCGGAAGATGTATCTGCCCTAGGAAATCATTTCACTGA 37743
1761 uSerGlyLeuValGlyLeuLeuGluGlnAlaAlaAspPheSerMet 1778
|:::| |||||:::|:|:|:|:|:|:|:|:|:|:|:|
37742 AACCTGGTTGAAGCGCTTGTGGAAGAAGCCTCCAAATCTTTTCAAGTTG 37693
1778 laGlyMetTyrGluAlaValAsnGluValTyrLysValLeuIleProIle 1794
|||||:::| |||||:::|:|:|:|:|:|:| |||||
37692 CTGGCATGTATGAAGCAATGAACGAAGTGTACAAATCTAATACCATTA 37643
1795 HisGluAlaAsnArgAspAlaLysLysLeuSerThrIleHisGlyLysLe 1811
|||||:::| |||||:::|:|:|:|:|:|:|:|:|:|:|:|
37642 TGGCAGGCTTAACAGAGATTTTCAAAAGCTAAGCAAGTTCATGGCAAAAT 37593
1811 uGlnGluAlaPheSerLysIle..... 1818
|||||:::| |||||:::|:|:|:|:|:|:| |||||
37592 GCAGGAGGCATTTAATCGAATATCCCAACTACAGTAACAATATTGTGTA 37543
1819ValHisGlnSerThrGlyTyrGluArgMet 1828
37542 AATTTTACCACGGAATATATACATATTATTAACAGCGTAAAGAGAGTT 37493
1829 PheGlyThrTyrPheArgValGlyPheTyrGlyThrLysPheGlyAspLe 1845
|||||:::| |||||:::|:|:|:|:|:|:| |||||
37492 TTTGGAACATACTTTCGTGTGGCTTCTATGCGCGAANAATTGGGCACTT 37443
1845 uAspGluGlnGluPheValTyrLysGluProAlaIleThrLysLeuAlaG 1862
|||||:::| |||||:::|:|:|:|:|:|:| |||||
37442 GGAATCAGCAGGAATTCATTTATAAAGACCAACATTGACGAAGTTGCCCG 37393
1862 lueSerHisArgLeu.Glu..... 1868
37392 AAATATTTAGTCGCTTCAGGTATATATTGCCAAATTGGAANAATAGAAC 37343
1869GlyPheTyrGlyGluArgPheG 1876
37342 TAATCAATTTGTTTCAACATACGTTAGAACCTTTTACACTGAACGATTCG 37293
1876 lylGluAspValValGluValIleLysAspSerAsnProValAspLysCys 1892
|| ||| |||||:::|:|:|:|:|:|:| |||||
37292 GACCGGACTCTGTGCATATCATTTAAAGATTCCAATACCGTTGATATTAAT 37243
1893 LysLeuAspProAsnLysAlaTyrIleGlnIleThrTyrValGluProTy 1909
:::| |||||:::|:|:|:|:|:|:| |||||
37242 AGCTTGATCCCGATAGGCTTACATTCAAATTACTTATGTTGAACCTTA 37193
1909 rPheAspThrTyrGluMetLysAspArgIleThrTyrPheAspLysAsnT 1926
|||||:::| |||||:::|:|:|:|:|:|:| |||||
37192 CTTTGAACATATGAATGCGTCATCGTGAGACATACTTTGAGCGGAATT 37143
1926 yrAsnLeu..... 1928
37142 TCAATATAAGTATGATATGAATTAAACAGATAATTAAATCGAAATTTAA 37093
1929Arg.ArgPheMetTyrCyst 1935
37092 AATTATAATTTTAACATTTTGTGTTATGCGAAGCGTTTATATATGCCA 37043
1935 hrProPheThrLeuAspGlyArgAlaHisGlyGluLeuHisGluGlnPhe 1951
|||||:::| |||||:::|:|:|:|:|:|:| |||||
37042 CGCCTTTTACTAAAAATGGAAGGCACATGGCAACTAAATGAGCAGTGT 36993
1952 LysArgLysThrIleLeuThrThrSerHisAlaPheProTyrIleLysTh 1968
|||||:::| |||||:::|:|:|:|:|:|:| |||||
36992 AAACGAAACGATTTTGACACAGCGCAATTCATTTCCATACGTAANAAC 36943
1968 rArgValAsnValThrHisLysGluGluIleIleLeuThrProIleGluV 1985
|||||:::| |||||:::|:|:|:|:|:|:| |||||

36942 TCGCATATGTAATTAGCCGACAGCAATAAGTACCTGGAACCGATTGAAG 36893
1985 alAlaIleGlu..... 1988
|||||||
36892 TGGCAATTGAAGTAAGTTATGTGTAAATCCCAAAAGTTAATAAATTGA 36843
1989AspMetGlnLysThrGlnGluLeuAlaPheAl 2000
|||||||
36842 AACGTTATTGCCATGATATTCAAAAAAACATTTGGAATTGGCAGCTGC 36793
2000 aThrHisGlnAspProAlaAspProLysMetLeuGlnMetValLeuGln 2017
|||||||
36792 CACGAATCAAGAGCCAGACAGACCCCAAATATTGCAAAATGTTACTGCAGG 36743
2017 lYserValGlyThrThrValAsnGlnGlyProLeuGluValAlaGlnVal 2033
|| :|||
36742 GATGTAATGGAACCACTGTTAACCAAGACCGATGGAATGGCGAGTGTG 36693
2034 PheLeuSerGluIleProSerAspProLysLeu..PheArgHisHisAs 2049
|||||||
36692 TTCCTTTCCAATTATTCCGACGAGAACAACTGTACCGACAAACACCAAAA 36643
2049 nLysLeuArgLeuCySPheLysAspPheThrLysArgCysGluAspAlaL 2066
|||||||
36642 CAAACTTCGGTGTGCTTTCGCGAGTTTCAAGCGTTGTGCTGATGCTT 36593
2066 euArgLysAsnLysSerLeuIleGlyProValGlnLysGluTyrGlnArg 2082
||:|||||
36592 TGAAGAAAGAAATCGCAATCTAATACTTTCAAGATCAAAAAAGATTACCAACGA 36543

2083 GluLeu 2084
|||||
36542 GAACGTG 36537

seq_name: gb_in:AC008318

seq_documentation_block: 166626 bp DNA linear INV 05-APR-2001
LOCUS AC008318
DEFINITION Drosophila melanogaster, chromosome 2L, region 21A-21A, BAC clone
BACR09J20, complete sequence.

ACCESSION AC008318
VERSION AC008318.7 GI:13549312
KEYWORDS HTG.

SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

REFERENCE 1 (bases 1 to 166626)
AUTHORS Celniker,S.E., Adams,M.D., Kronmiller,B., Tyler,D., Wan,K.H.,
Holt,R.A., Evans,C.A., Gocayne,J.D., Amanatides,P.G., Brandon,R.C.,
Rogers,Y., An,H., Baldwin,D., Banzon,J., Beeson,K.Y., Busam,D.A.,
Carlson,J.W., Center,A., Champe,M., Davenport,L.B., Dietz,S.M.,
Dodson,K., Dorselt,V., Doup,L.E., Doyle,C., Dresnek,D., Farfan,D.,
Ferriera,S., Frise,E., Galle,R.F., Garg,N.S., George,R.A.,
Gonzalez,M., Houck,J., Hoskins,R.A., Hostin,D., Howland,T.J.,
Ibegwam,C., Jalali,M., Kruse,D., Li,P., Mattei,B., Moshrefi,A.,
McIntosh,T.C., Moy,M., Murphy,B., Nelson,C., Nelson,K.A., Nunco,J.,
Pacleb,J., Paragtas,V., Park,S., Patel,S., Pfeiffer,B.,
Phouanavong,S., Pittman,G.S., Puri,V., Richards,S., Scheeler,F.,
Stapleton,M., Strong,R., Svirskas,R., Tector,C., Williams,S.M.,
Zaveri,J.S., Smith,H.O., Rubin,G.M. and Venter,J.C.

TITLE Sequencing of Drosophila chromosome 2L, region 21A-21A

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 166626)

AUTHORS

Celniker,S.E., Agbayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G.,
Butenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L.,
Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L.,
Hoskins,R.A., Houston,K.A., Humastli,S.R., Karra,K., Kearney,L.,
Kim,E., Lee,B., Lewis,S., Li,P., Lomotan,M.A., Mazda,P.,
Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S.,
Pfeiffer,B., Poon,L., Sequeira,A., Sethi,H., Snir,E.,
Svirskas,R.R., Wan,K.H., Weinburg,T., Zhang,R., Zieran,L.L. and

Rubin,G.M.
TITLE Direct Submission
JOURNAL Submitted (02-AUG-1999) Drosophila Genome Center, Lawrence Berkeley
Laboratory, MS 64-121, Berkeley, CA 94720, USA
COMMENT On Apr 5, 2001 this sequence version replaced gi:6728945.
Sequence submitted by:
Berkeley Drosophila Genome Project
Lawrence Berkeley National Laboratory, MS 64-121
Berkeley, CA 94720

This sequence was assembled using end sequences from a whole genome
shotgun and from subclones of this BAC and its neighboring clones.
For further information about this sequence, including its location
and relationship to other sequences, please visit our sequence
archive Web site (<http://www.fruitfly.org/sequence/>) or send email
to bdg@fruitfly.berkeley.edu.

FEATURES
source location/Qualifiers
1. 166626
/organism="Drosophila melanogaster"
/strain="y; cn bw sp"
/db_xref="taxon:7227"
/chromosome="2L"
/map="21A-21A"
/clone="BACR09J20 (D841)"
/clone_lib="RPCI-98 (Roswell Park Cancer Institute
Drosophila melanogaster BAC library, partial EcoRI in
pBACe3.6)"

BASE COUNT 49737 a 32951 c 34017 g 49921 t
ORIGIN

alignment_scores: Quality: 4690.00 Length: 2340
Ratio: 2.886 Gaps: 58
Percent Similarity: 69.444 Percent Identity: 44.188

alignment_block:

US-09-737-246-2 x AC008318 ..

Align seg 1/1 to: AC008318 from: 1 to: 166626

4 ArgArgAlaPheAlaGlnLysIleSer..... 12
:::|||||
85506 CAGCGCTCGTTGTGCAAAAAGTTTCCAAGTAAGTTTTCACCTAAAGGAA 85555
12 12
85556 ATAGTTATTTTAAACAAGAAGCGCCCTTAATGCTGTATACATATGTA 85605
13ArgThrValAlaAlaGluValArgLys 21
||| |||||
85606 CATATATCAATTGTGAATTCCTAGCGCAGTCGGCTGCGGATGACGAAAA 85655
22 GlnIleSerGlyGlnTyrSerGlySerProGlnLeuLysAsnLeuAs 38
:::||||| ::| ::| ::|
85656 AATGTTTCTGTTGCCATCTATCAAAAGGCAATG.....GACCCATC 85696
38 nIleValGlyAsnIleSer.HisHisThrThrValProLeuThr..... 52
::: ||||| ::| ::|
85697 ATTGTGTGTTCTTCGTTAGAGACCTATTGTTTAAATATATCAATTTT 85746
53GluAlaValAspProValAs 59
85747 CTGTTAAGCATTCATTTTTCATATAGATCTCCCGCAGACGCCCTTGA 85796
59 pLeuGluAspTyrLeuIleThrHisProLeuAlaValAspSerglyProl 76
| ||||| ||| ::| ::|
85797 TTACGAAGAATTCTATCCAGCATATGAACATCATCAACCGGATCCCT 85846
76 euArgAspLeuIleGluPheProProAspAspIleGluValValTyrSer 92
||:| ::| ::| |||
85847 TGAACATATATATTGGACTTTCCTCAAGGGGATGTAAATCATT 85896
93 ProArgAspCysArgThr.LeuValSerAlaValProGlu..... 105

85897 CCCCCAAAATCAGAACTGTTGATCACGTTGTTCCTCAATGAGAAATTTGTA 85946
106Glu 106
85947 TGTATGCACCTTAAGTATATACACAGCTACATAGCTACAGTCTCTTTTAGG 85996
107 SerGluMetAspProHisValArgSpCysIleArgSerTyrThrGluAs 123
85997 TCTGATTTGCCCATCAGCTACAGAAATGTGTAACTGCTATACGCGCCC 86046
123 pTrpAlaIleVal.....I 128
86047 ATGGAAGGTGGTTGAATATGCACAGCGCACCTTTCCAGTTCTTGCTATA 86096
128 leArgLysTyrHisLysLeuGlyThrGlyPheAsnProAsnThrLeuAsp 144
86097 TCCGCGAGCGTATCGACCGCGGAACC...ATAAGTCCCTCGCTACACAG 86143
145 LysGlnLysGluArgGlnLysGlyLeuProLysGlnValPheGluSerAs 161
86144 CAGGAATTCGAGATCGACACAAGGAT.....TTTTCGTCAATT 86178
161 pGluAlaIleProAspGlyAsnSerTyrGlnAspSpGln..AspAspLeuL 177
86179 TGA.....GAGCGCTCGCTTATATAAAGCGAGAGCTGCACCTCAAGTT 86222
177 ysArgArgSerMet.....SerIleAspasp... 185
86223 CTGACAAATCCATTCGAAGTTGGCTTCAGTAAGCTCTTGACGACACT 86272
186 ...ThrProArgGlySerTrpAlaCysSerIlePheAspLeuLysAsnSe 201
86273 TTGACCCCGCGTGGTCTTGCGGTAGC.....TTTGATCTTCGACGTTTC 86316
201 rLeuProAspAlaLeuLeuProAsnLeuLeuAspArgThrProAsnGluG 218
86317 GGTTAATGACCCACTAATACCCAAATTGCTTGACAATGTACTCCGAGAC 86366
218 luIleAspArgGlnAsnAspArgGlnArgLysSerAsnArgHisLysGlu 234
86367 ATATGATCAATCCAATGAGACAGCGCGCAGACAAGATCGACAATGGCT 86416
235 LeuPheAlaLeuHisProSerProAspGluGluProIleGluArgLe 251
86417 CTTTCTCACTGTATCCGGAATCTGAACCGGAGGAAACATTTGAACGCCG 86466
251 uSerValProAspIleProLysGlnHisPheGlyGlnArgLeuLeuVal 268
86467 CCTGCTCGCAGAGATTCCAGTTGAACATATGGTCAATCGCATTCAGGTTA 86516
268 ysCysLeuSerLeuLysPheGluIleGluIleGluProIlePheAlaSer 284
86517 ACTGCTTACAGCTTCGGCTGAGCTGGAAGTGAGCCAAATTTTGCATCG 86566
285 leuAlaLeuTyrAspValLysGluLysLysIleSerGluAsnPheTy 301
86567 ATGGCTATTATGATGCTAAGGAACGGCAAAAAATTCTGAATAATTCTTA 86616
301 rPheAspLeuAsnSerGluGlnMetLysGlyLeuLeuArgProHisValP 318
86617 TTTTGACATGAACGAGATAGTCTTAACGCATGTGTCAAGTCATGTAC 86666
318 rOPrOAlaAlaIleThrThrLeuAlaArgSerAlaIlePheSerIleThr 334
86667 AGTGGCGCGACATTAAGTACCCAGAGCCACTCAGCCATATTGCAAAATTTCA 86716
335 TyrProSerGlnAspValPheLeuValIleLysLeuGluLysValLeuG 351
86717 TATCGAGTAACGATTTATTTCTGATTCGTGTGGAATAAGTTTGTG.. 86764
351 ngInGlyAspIleGlyLysAlaGluProTyrMetIlePheLysGluA 368

86765 .CAGGGTGACATAACAAATTCCGTCGAGCCGTATCTC..... 86800
368 laAspAlaThrLysAsnLysGluLysLeu...GluLysLeuLysSerGln 383
86801AAGAAGACAAAGATAAATACCAGAGAGAAAGTAAATCCAAC 86842
384 AlaAspGlnPheCysGlnArgLeuGlyLysTyrArgMetProPheAlaTr 400
86843 GCTGTGAGCTACTGTGAACGAGCTAGGCAAAATACCGAATGCGTTGCGATG 86892
400 pThrAlaIleHisLeuMetAsnIleValSer..... 410
86893 GACAGGAATCTACCTAACCAAAATGTTTAAATGGCGATTAACCTTGAAAGCA 86942
411 ..SerAlaGlySerLeuGluArgAspSer..... 419
86943 AAGACGCTGGCGCCCGGGAACGGGATTCCTTTGGAATGCTTGTGGCAGC 86992
420ThrGluValGluIleSerThrGlyGluArgLysGlySerTr 433
86993 AGCTTGGGCACAGCGCCCAAGCTCCCATAGCTTAGATCGAAATCTTCCAC 87042
433 pSerGlu.....ArgArg.....AsnSerSerI 441
87043 AAGTAGCTTTGATCACTAAGGCGGAAAGCTTAACGATATGAGCGGACAC 87092
441 leValGlyArgArgSerLeuGlu.....ArgThrThrSer 452
87093 TTACACGACGCTGCTTCCTTGAAACGGAAGAAAGCACCGATCCTGGTCC 87142
453 GlyAspAspAlaCysAsnLeuThr..SerPheArgProAlaThrLeuTh 468
87143 CCAGATGATTTTGCAAAAGTTGTAGAAACTTTCGCTACACTATATAAC 87192
468 rValThrAsnPhePheLysGlnGluGlyAspArgLeuSerAspGluAspL 485
87193 AGTACCAAGCTTCTTTAAACAAGAGGAGATAAATGAAGAGCAGGATT 87242
485 euTyrLysPheLeuAlaAspMetArgArgProSerSerValLeuArgArg 501
87243 TGTACAAATAATTTACCAGAGAGTTAAACGACCGAGCTCAGTTATGAGAAG 87292
502 leuArgProIleThrAlaGlnLeuLysIleAspIleSerProAlaProGl 518
87293 TACAATGCAATACCTGGTCTATTAATTAAGAATATCACCTTGGCTCGA 87342
518 uAsnProHisTyrCysLeuThrProGluLeuGlnValLysLeuTyrP 535
87343 GGAAGCCAGAATGCATTAACGCCAGAGTTAGCACCTATAAACCCCTCAA 87392
535 roAspSerArgValArgProThrArgGluIleLeuGluPheProAlaArg 551
87393 GTGCAGATTAATGTGCGCCAGTTAAGCAAACTTAGAATTTCCGCAATCG 87442
552 AspValTyrValProAsnThrThrTyrArgAsnLeuLeuTyrIleTyrP 568
87443 GCGATTTACAATCCACATTTATAGCTATGCCAATCTCCTATTGTGTCACC 87492
568 ogInSerLeuAsnPheAlaAsnArgGlnGlySerAlaArgAsnIleThrV 585
87493 AAAGGAGCTAATTTTCTTCGCGTGCAGGTTCTGCTCGAACAATTCGCCG 87542
585 alLysValGlnPheMetTyrGlyLysProSerAsnAlaMetProVal 601
87543 TTGCGGTTCACTAATATGGCTGGAGAAACCCCAAAAGATGCGGTCAATGCC 87592
602 IlePheGlyLysSerSerCysSerGluPheSerLysGluAlaTyrThrAl 618
87593 ATCTACGGCAAGTCTGTTGTGCCAAATTTTCTACTGAAGCAATTACAGC 87642
618 aValValTyrHisAsnArgSerProAspPheHisGluIleLysValL 635
87643 TGTCAATATATCATAACAAAATGCCCATTCATTCTACGATGAATAAAAATTG 87692

635 ysleuProAlaThrLeuThrasPHisHisIleuLeuPheThrPheTyr 651
|||||::: ::|||
87693 CGCTGCCTGCATCAATAAGACGATCACCATTATTTGTCACCATTTAC 87742
652 HisValSerCysGlnGlnLysGln.....AsnThrProLeuGluTh 665
|||||::: ::|||
87743 CATGTTTCATGTCAAAGAAACACACAGACTTGCAGCCGCTCAGTAGAGAC 87792
665 rProValGlyTyrThrTrpIlePrometLeuGlnAsnGlyArgLeuLysT 682
|||||::: ::|||
87793 TCCTATCGGCTACACATGCTGCCCTTACTGGAAGATGGAACCTTAAGT 87842
682 hrGlyGlnPheCysLeuProValSerLeuGluLysProGlnAlaTyr 698
|||||::: ::|||
87843 TTGGGAATTTAATCTTCCTGTTATGTTAGTAATCACCGCGAAATTTAC 87892
699 SerValLeuSerProGluValProLeuProGlyMetLysTrpValAspAs 715
|||::: |||::: |||
87893 TCATTTATACCAACCGAATGTTCACTGCTGGAATTAATGGCTGACAA 87942
715 nHisLysGlyValPheAsnValGluValAlaValSerSerIleHisT 732
|||||::: ::|||
87943 TCACAGAGCCGCTGTTTCATTAATGTAGAGCAGTCACCGCAATTCATA 87992
732 hrGlnAspProTyrLeuAspLysPhePheAlaLeuValAsnAlaLeuAsp 748
|||::: |||::: |||
87993 CTTTGGATTCTTCCCTGCATCGATTCTTCTTAATTTGCCGAATATCTAGAC 88042
749 GluHisLeuPheProValArgIleGlyAspMetArgIleMetGluAsnAs 765
::: ::||| ::|||
88043 ACTCGAAATATACCTTCCCATTTGGC.....GAGAAATAA 88077
765 nLeuGluAsnGluLeuLysSerSerIleSerAlaLeuAsnSerSerGlnL 782
|||::: |||::: |||
88078 TATAGAAACAGAAATTAAAGAAATGCTTGCTGATATCGAATATGCAAAATC 88127
782 euGluProValValArgPheLeuHisLeuLeuAspLysLeuIleLeu 798
|||||::: |||
88128 GTAGCCATTAGTAGAGCATTACCATTAAGTCTGATTAAGCTTATCGAA 88177
799 LeuValIleArgProProValIleAlaGlyGlnIleValAsnLeuGlyG 815
|||::: |||
88178 TTGTTAGTTGTCACCCACAAGTTGGTGACAAAGCAATGCTCTGGGATC 88227
815 nAlaSerPheGluAlaMetAlaSerIleIleAsnArgLeuHisLysAsnL 832
::: |||
88228 TACGGTTTTCGAAGTCTTGTGTTGGTATCATCTTGTATCATACTTAA 88277
832 euGluGlyAsnHisAspGlnHisGlyArgAsnSerLeuLeuAlaSerTyr 848
::: |||
88278 ACGACGATCAATATGACCAATACGGTGGCAAGTTTGTATCAACATAT 88327
849 IleHisTyrValPheArgLeuProAsnThrTyrProAsnSerSerSerPr 865
::: |||
88328 GTGCACCTTCAATGTAAATTCGCGACCCCTTCAA..... 88363
865 oGlyProGlyGlyLeuGlySerValHisTyrAlaThrMetAlaArgS 882
||| ::|||
88364AGCAACAGCGACTAACATGCAAGTCGCA 88391
882 eraAlaValArgProAlaSerLeuAsnLeuAsnArgSerArgSerLeuSer 898
|||::: |||
88392 GCACACAGCAA.....GACTTGCATTTAAGCGAATCGTACACCATATAT 88435
899 AsnSerAsnProAspIleSerGlyThrProThrSerProAspAspGluVa 915
::: |||
88436 GATAAT..... 88441
915 lArgSerIleIleGlySerLysGlyLeuAspArgSerAsnSerTrpValA 932
::: |||
88442 ...GTTTACCAAGTGCGCGAAGCTTGATCGCAAAAGTTAA.....A 88480

932 snThrGlyGlyProLysAlaAlaProTrpGlySerAsnProSerProSer 948
|||||
88481 ATACATATTTTGTAAAGGCT..... 88500
949 AlaGluSerThrGlnAlaMetAspArgSerCysAsnArgMetSerSerHi 965
|||
88501TTAAATATTAACCTTTTCCT 88520
965 strhGluThrSer...SerPheLeuGlnThrLeuThrGlyArgLeuProT 981
::: |||
88521 TTCAGAGCTGTCCATAGATGTCTTCACTCGATGGCTGCACGGATGTCC 88570
981 hrLysLysLeuPheHisGluGluLeuAlaLeuGlnTrpValAlaCysSer 997
::: |||
88571 AAGTGGGCTGCTTCACGAAAGATTAGCCTTGCAATGGGTGTTGCAAGT 88620
998 GlySerValArgGluSerAlaLeuGlnGlnAlaTrpPhePheGluLe 1014
|||::: |||
88621 GGAAGGCAAGCTGACTTAGCCATGTCAAAATTCGTGCTTTTATTTGAGCT 88670
1014 uMetValLysSerMetValHisHisLeuTyrPheAsnAspLysLeuGluA 1031
|||::: |||
88671 TATGTCAAAATCTATGATTGAACATTTGCAATGTTCAAAATCTTAAATG 88720
1031 lAProArgLysSerArgPheProGluArgPheMetAspAspIleAlaAla 1047
::: |||
88721 GTCCCGCAACATCGATTCTCCACATCAGTTTAATGATGATATATCCACC 88770
1048 LeuValSerThrIleAlaSerAspIleValSerArgPheGlnLysAspTh 1064
|||||
88771 CTGGTGCACTTAGTTACACAAAGGTGTCGGCTACACACAGCAACGAACC 88820
1064 rGluMetValGluArgLeuAsnThrSerLeuAlaPhePheLeuAsnAspL 1081
::: |||
88821 TAAATTAGCGCAGTCATTAAATGCACAGTTTAGGCTTTTATATTGACA 88870
1081 euLeuSerValMetAspArgGlyPheValPheSerLeuIleLysSerCys 1097
::: |||
88871 TACTAAGCATCATGGATCGGGGATTTGTTTCGGCCTCATCAAAACGTAC 88920
1098 TyrLysGlnValSerSerLysLeuTyrSerLeuProAsnProSerValLe 1114
|||
88921 ACAAAAGTTCGATTCTTAAAAATGCTTCAATACCGGAT.....CT 88961
1114 uValSerLeuArgLeuAspPheLeuArgIleIleCysSerHisGluHisT 1131
|||
88962 AATGAACATAAATAAGACTTTTGTGAAATGTGTGTAGTCAATGAGCATT 89011
1131 yRValThrLeuAsnLeuPro.....CysSerLeuLeuThrPro 1143
::: |||
89012 TTGTAGCCTTAAACTTGGCGTTTGGTACTTCATACACAATGGTAAACAGCG 89061
1144 ProAlaSerProSerProSerValSerSerAlaThrSerGlnSerSerG 1160
|||
89062 CCATGCAGCCCCCAGCCCAAGCACACGTCGACACAGCCAAACTTCTTG 89111
1160 yPheSerThrAsnVal..... 1165
|||
89112 CGTAAGTAAAAAACAATACACTTTAATTTTACTTATTGTTGGCTTCTT 89161
1166GlnAspGln..LysIleAlaAsnMetPheGluLeuSerValProp 1180
::: |||
89162 TTAAAGGGATCATAGAGAGAGCCCTTACGCCGACCTTAGTCAAGAAT 89211
1180 heArgGlnGlnHisTyrLeuAlaGlyLeuValLeuThrGluLeuAlaVal 1196
|||
89212 TCCTCCAGCAACATTTTCTAGTCGAGACTTGTCTGAGCGACTTAGCAGCA 89261
1197 IleLeuAsp...ProAspAlaGluGlyLeuPheGlyLeuHisLysLysVa 1212
::: |||
89262 GTAATGAGGTGCCAACCACCAA.....TTACATGGAAAGC 89299
1212 IleAsnMetValHisAsnLeuLeuSerSerHisAspSerAspProArgT 1229

89300 AATAAGGTGTATTCCGAACCTGATGACATCTCACGATTTTGACCGCGGT 89349
1229 ySerAspProGlnIleLysAlaArgValAlaMetLeuTyrLeuProLeu 1245
89350 ATAGCGAAACTGATGCACGTCGAAGGTGGCTGCTTATATATACCGCTA 89399
1246 IleGlyIleIleMetGluThrValProGlnLeuTyrAspPhe..... 1259
89400 CTTTCTATAGTGTGATGATGATTCCTCAGTTGCATCAGCACGCGCTTGA 89449
1260 .ThrGluThrHisAsnGlnArgGlyArgProIleCysIleAlaThrAspA 1276
89450 GCAGATCGTCTGCACGCAAAATCGGTCAA.....CTTGAAG 89484
1276 sPTYrGluSerGlu.....SerGlySerMetIleSerGlnThr 1288
89485 ACTATCAAGGGCCACATCAAACTATACAGCATCAACCATAAATCCGGAA 89534
1289 ValAlaMetAlaIleAlaGlyThrSerValProGlnLeuThrArgProG 1305
89535 GTGGCATTTGGCATATCGGGA.....AGTCGTCCCTA 89566
1305 ySerPheLeuLeuThrSerThrSerGlyArgGlnHisThrThrPheSerA 1322
89567 CTCCTACTTG....AATGAACAAGTTAAAAACAATTCGCCCTCAGCT 89610
1322 IaGluSerSerArgSerLeuLeuIleCysLeuLeuTyrValleuLysAsn 1338
89611 CGGGAATACCCCGCATTTATGTGTGTTTCTGTGGTCTTAAAAAAC 89660
1339 AlaAspGluThrValleuGlnLysTrpPheThrAspLeuSerValleuGl 1355
89661 TTGAGCGCAAGTTTATATCGTGGCTTATGGATTTAAGTCCGCATAG 89710
1355 nLeuAsnArgLeuLeuAspLeuLeuTyrLeuCysValSerCysPheGluT 1372
89711 AGTGCACCAAAATGCTTCAAGGTGTTAACGTTTGTGAACCTTTCGAGT 89760
1372 yTrLysGlyLysLys.....ValPheGluArgMetAsnSerLeuThr 1385
89761 ACACAGACAAAAAAACGTCGCCACCCTTAAACGACGAATACACAAAGT 89810
1386 PheLysLys.....SerLysAspMetArgAlaLysLeuGluAlaIle 1400
89811 TTTAGGAAAACTGGATCTACTGATGTTAAGGAAAAGCTGGAAGATGTAT 89860
1400 eLeuGlySerIleGlyAlaArgGlnGluMetValArgArgSerArgGlyG 1417
89861 TAGAGGCACAAATTTCAGCACGTTATGATTTAATTAATCGTCAAAAAGATC 89910
1417 InLeuGluArgSerProSerGlySerAlaPheGlySerGlnLysAsnLeu 1433
89911 GT.....AATTCAACAGAAAAAATTTC 89930
1434 ArgTrpArgLysAspMetThrHisTrpArg..GlnAsnThrGluLysLe 1449
89931 CGATGGAGGAAAGACCAGATGCCATATCGCTCTCAATATGCCGATGCTGT 89980
1449 uAspLysSerArgAlaGluIleGlnHisGluAlaLeuIleAspGlyAsnL 1466
89981 GGGAAAAAGTGAGCATGAACTTGAATTAAATCATTTTCGAAAGGCTCTT 90030
1466 euAlaThrGluAlaAsnLeuIleIleLeuAspThrLeuGluIleValVal 1482
90031 TGGCAACTGAAGTTGCACCTTGACTTCTCGATACCCCTTGAATTAATTGTT 90080
1483 GlnThrValSerValThrGluSerLysGluSerIleLeuGlyGlyValLe 1499
90081 CAT.....GCTGCTGCCAACCTTATACAAATCTTCTTGAACCTGTGCT 90124
1499 uLysValLeuLeuHisSerMetAlaCysAsnGlnSerAlaValTyrLeuG 1516

90125 AAAGGTGCTTCTTCATCTTTGTGTCGCCGAATCAATCTGTTTGGCCCTGC 90174
1516 InHisCysPheAlaThrGlnArgAlaLeuValSerLysPheProGluLeu 1532
90175 AAAACTGTTTGGCTCCCAACGCGCGTTAATCTTTAAATTTCCAAACTTA 90224
1533 LeuPheGluGluGluThrGluGlnCysAlaAspLeuCysLeuArgLeuLe 1549
90225 TTATTTGATGACGAGACCGATATATGTGCGGATTTGTCTAATTTTACT 90274
1549 uArgHisCysSerSerSerIleGlyThrIleArgSerHisProSerAlas 1566
90275 GAAACATTTGGGTCCCTATTGCCAGGAATGACGATCACAGCAGCTGCAT 90324
1566 erLeuTyrLeuLeuMetArgGlnAsnPheGluIleGlyAsn..... 1579
90325 CACTATATTACTAATGAGACAAAAATTTGAAATTTGGAATGTAGTTT 90374
1579 1579
90375 ATTTTATTGTTTACGTTTATTTAACTAATTAATGCTACTTATTGCTAT 90424
1580 .Asn.PheAlaArgValLysMetGlnValPrometSerLeuSerSerLeu 1595
90425 AGAACTTTGCCCGAGTTAAGATGCAAGTGACGATGCTTTAAGCTCCCTA 90474
1596 ValGlyThrSerGlnAsnPheAsnGluGluPheLeuArgArgSerLeuLy 1612
90475 GTTGAACAAGTTCGGTCTTTAGTGAGCAATCCTTACGCCGTGCCGTAA 90524
1612 sThrIleLeuThrTyrAlaGluGluAspLeuGluLeuArgGluThrThrP 1629
90525 AACCGTCCGTGTTTATGCTGAATCCGACTCCGACCTGCAGGATCTTCTT 90574
1629 heProAspGlnValGlnAspLeuValPheAsnLeuHisMetIleLeuSer 1645
90575 TTCCTGAACAAGTACAAAGATTGCTTTTAATCTGCATATGATCCTGTCA 90624
1646 AspThrValLysMetLysGluHisGlnGluAspProGluMetLeuIleAs 1662
90625 GATACTGTTAAATGAAGAGTATCAGGAAGACCCAGAATGTTCCTTGA 90674
1662 pleuMetTyrArgIleAlaLysGlyTyrGlnThrSerProGlu...ArgL 1678
90675 CCTCATGAATCGTATTGCCAAGGAGTACCAAAATAACCTCATCTACGAC 90724
1678 euThrTrpLeuGlnAsnMetAlaGlyLysHisSerGluArgSerAsnHis 1694
90725 TGACTTGGTTGAAAATATGGCTAAAAAACACCGCGAGCGAGCAAAATCAC 90774
1695 AlaGluAlaGlnCysLeuValHisSerAlaAlaLeuValaGluTyr 1711
90775 ACGGAAGCAGCCATGTTTATGTACATGCTGCTTCTTAGTTCTGAATA 90824
1711 rLeuSerMetLeuGluAspArgLysTyrLeuProValGlyCysValThrp 1728
90825 TCTTAGCATGTTGGAGTCACAAAAACATTTGCCCTGTTGGAGCTGTAAGTT 90874
1728 heGlnAsnIleSerSerAsnValLeuGluGluSerAlaValSerAspAsp 1744
90875 TTCACGCAATTTCTCCCAACACACTAATGAGTGGCGCGGTATCGGATGAT 90924
1745 ValValSerProAspGluGluGlyIleCysSerGlyLysTyrPheThrGl 1761
90925 GTGCTAGTCCCGCGGGAAGATGTATCTGCCTAGGAATCATTTTCACTGA 90974
1761 uSerGlyLeuValGlyLeuLeuGluGlnAlaAlaAlaSerPheSerMeta 1778
90975 AACTGGGTTGAAGCCTTGTGGAAGAACCTCCAATTCCTTTCAAGTTG 91024
1778 IaGlyMetTyrGluAlaValAsnGluValTyrLysValleuIleProIle 1794
91025 CTGGCATGTATGAAGCAATGAAGCAAGTGTACAAAATCTAATACCATATA 91074

1795 HisGluAlaAsnArgAspAlaLysLysLeuSerThrIleHisGlyLysLeu 1811
|||||:::|||||:::|||||:::|||||:::|||||
91075 TGGCAGGCTTAACAGAGATTTCAAAAGCTAAGCAAGTTCATGGCAATTT 91124
1811 uGlnGluAlaPheSer..... 1816
|||||:::|||||:::|||||:::|||||:::|||||
91125 GCAGGAGGCATTTAATCGAATATCCCACTACAGGTACATATTTGTGTA 91174
1817LysIle.ValHisGlnSerThrGlyTrpGluArgMe 1828
|||||:::|||||:::|||||:::|||||:::|||||
91175 AATTTTACCACGAGAAATATATACATATTTATAACAGGGTAAAGAGAGT 91224
1828 tPheGlyThrTyrPheArgValGlyPheTyrGlyThrLysPheGlyAspL 1845
:|||||:::|||||:::|||||:::|||||:::|||||
91225 TTTTGGACATACCTTTCGTGGCTTCTATGGCGAAATTTGGGGACT 91274
1845 euAspGlnGlnGluPheValTyrLysGluProAlaIleThrLysLeuAla 1861
|||||:::|||||:::|||||:::|||||:::|||||
91275 TGGATCAGCAGGAATTCATTTATAAGACCAACATTCAGCAAGTTGCCCC 91324
1862 GluIleSerHisArgLeu.Glu..... 1868
|||||:::|||||:::|||||:::|||||:::|||||
91325 GAAATATTTAGTCGGCTTCAGGTATATATTGCAAAATGGAAAAATAGAA 91374
1869GlyPheTyrGlyGluArgPhe 1875
:|||||:::|||||:::|||||:::|||||:::|||||
91375 CTAATCAATTTTGTTCACACATACGTTAGAACTTTTACACTGAAACGATTC 91424
1876 GlyGluAspValValGluValIleLysAspSerAsnProValAspLysCy 1892
||| ||| |||:::|||||:::|||||:::|||||
91425 GGACCGGACTCTGTGCATATCATTAAGATTCCAATACCGTTGATATTAA 91474
1892 sLysLeuAspProAsnLysAlaTyrIleGlnIleThrTyrValGluProT 1909
:::|||||:::|||||:::|||||:::|||||:::|||||
91475 TAGCTTGGATCCCGCTAAGGCTTACATTCAAATTACTTATGTTGAACCCCT 91524
1909 yrPheAspThrTyrGluMetLysAspArgIleThrTyrPheAspLysAsn 1925
|||||:::|||||:::|||||:::|||||:::|||||
91525 ACTTTGAACATATGAATGCGTCCGTGAGACATACCTTGCAGCGGAAT 91574
1926 TyrAsnLeu..... 1928
:::|||||:::|||||:::|||||:::|||||
91575 TTCATATATAGTATGATATGAATTAACAGATAATTTAAATCGAAATTTA 91624
1929Arg.ArgPheMetTyrCys 1934
||| ||| |||:::|||||:::|||||:::|||||
91625 AAATTATATTTTAACATTTTGGTTTATGCGAAGCTTTTATATATATGCC 91674
1935 ThrProPheThrLeuAspGlyArgAlaHisGlyGluLeuHisGluGlnPh 1951
|||||:::|||||:::|||||:::|||||:::|||||
91675 ACGCCTTTACTAAAAATGGAAGGCACATGGCGAACTAAATGAGCAGTG 91724
1951 eLysArgLysThrIleLeuThrThrSerHisAlaPheProTyrIleLysT 1968
|||||:::|||||:::|||||:::|||||:::|||||
91725 TAAACGMAAACGATTTTGACCACGGCGAATCACTTTCATACGTAAAAA 91774
1968 hrArgValAsnValThrHisLysGluGluIleIleLeuThrProIleGlu 1984
|||||:::|||||:::|||||:::|||||:::|||||
91775 CTCGCATTTAGTAAATAGCCGACAGCAAAATAGCTCGAACCAGATTGAA 91824
1985 ValAlaIleGlu..... 1988
|||||:::|||||:::|||||:::|||||:::|||||
91825 GTGGCAATTTGAAGTATATGTGTAATAATCCCAAAAGTTAATAAATTT 91874
1989AspMetGlnLysLysThrGlnGluLeuAlaPhe 2000
|||:::|||||:::|||||:::|||||:::|||||
91875 AAACGTTATTTGCTAGGATATTCAAAAAACAATTGGAATTTGGCAGCTG 91924
2000 laThrHisGlnAspProAlaAspProLysMetLeuGlnMetValleuGln 2016
|||||:::|||||:::|||||:::|||||:::|||||
91925 CCACGAATCAAGAGCCAGACAGCCCAAAATATTTGCAAAATGGTACTGCAG 91974

2017 GlySerValGlyThrThrValAsnGlnGlyProLeuGluValAlaGlnVa 2033
||| :::|||||:::|||||:::|||||:::|||||
91975 GGATGTATTTGGAACCACTGTTAACCAAGACCGATGAAATGGCGAGTGT 92024
2033 lPheLeuSerGlnIleProSerAspProLysLeu..PheArgHisHisA 2049
|||||:::|||||:::|||||:::|||||:::|||||
92025 GTTCCTTCCAAATTATCCGACGGAACAACTGTACCGACAAACACCAAA 92074
2049 snLysLeuArgLeuCysPheLysAspPheThrLysArgCysGluAspAla 2065
|||||:::|||||:::|||||:::|||||:::|||||
92075 ACAAACTTCGTTGTGCTTTCGCGAGTTTCAAGCGTTGTGCTGATGCT 92124
2066 leuArgLysAsnLysSerLeuIleGlyProValGlnLysGluTyrGlnAr 2082
|||:::|||||:::|||||:::|||||:::|||||
92125 TTGAAAAAGAAATCGCAATCTAATACTTTCAGATCAAAAAGATTACCAACG 92174
2082 gGluLeu 2084
|||||
92175 AGAAGCTG 92181
seq_name: gb_in:AE003590

seq_documentation_block:

LOCUS AE003590 301639 bp DNA linear INV 04-OCT-2000
DEFINITION Drosophila melanogaster genomic scaffold 14200001386046 section 1
of 16, complete sequence.

ACCESSION AE003590 AE002638
VERSION AE003590.2 GI:10727416
KEYWORDS HTG.

SOURCE fruit fly.

ORGANISM Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

REFERENCE 1 (bases 1 to 301639)
AUTHORS

Adams,M.D., Celniker,S.E., Holt,R.A., Evans,C.A., Gocayne,J.D.,
Amanatides,P.G., Scherer,S.E., Li,P.W., Hoskins,R.A., Galle,R.F.,
George,R.A., Lewis,S.E., Richards,S., Ashburner,M., Henderson,S.N.,
Sutton,G.G., Wortman,J.R., Yandell,M.D., Zhang,Q., Chen,L.X.,
Brandon,R.C., Rogers,Y.H., Blazej,R.G., Champagne,M., Pfeiffer,B.D.,
Wan,K.H., Doyle,C., Baxter,E.G., Helt,G., Nelson,C.R., Gabor
Miklos,G.L., Abril,J.F., Aghayani,A., An,H.J.,
Andrews-Pfannkoch,C., Baldwin,D., Ballew,R.M., Beeson,A.,
Baxendale,J., Bayraktaroglu,L., Beasley,E.M., Beeson,K.Y.,
Benos,P.V., Berman,B.P., Bhandari,D., Bolshakov,S., Borkova,D.,
Botchan,M.R., Bouck,J., Broksstein,P., Brottler,P., Burtis,K.C.,
Busam,D.A., Butler,H., Cadien,E., Center,A., Chandra,I.,
Cherry,J.M., Cawley,S., Dahlke,C., Davenport,L.B., Davies,P., de
Pablos,B., Delcher,A., Deng,Z., Mays,A.D., Dew,I., Dietz,S.M.,
Dodson,K., Doup,L.E., Downes,M., Dugan-Rocha,S., Dunkov,B.C.,
Dunn,P., Durbin,K.J., Evangelista,C.C., Ferraz,C., Ferriera,S.,
Fleischmann,W., Fosler,C., Gabriellian,A.E., Garg,N.S.,
Gelbart,W.M., Glasser,K., Glodek,A., Gong,F., Gorrell,J.H., Gu,Z.,
Guan,P., Harris,M., Harris,N.L., Harvey,D., Helman,T.J.,
Hernandez,J.R., Houck,J., Hostin,D., Houston,K.A., Howland,T.J.,
Wei,M.H., Ibegwam,C., Jalali,M., Kalush,F., Karpen,G.H., Ke,Z.,
Kennison,J.A., Ketchum,K.A., Kimmel,B.E., Kodira,C.D., Kraft,C.,
Kravitz,S., Kulp,D., Lai,Z., Lasko,P., Lei,Y., Levitsky,A.A.,
Li,J., Li,Z., Liang,Y., Lin,X., Liu,X., Mattei,B., McIntosh,T.C.,
McLeod,M.P., McPherson,D., Merkulov,G., Milshina,N.V., Mobery,C.,
Morris,J., Moshrefi,A., Mount,S.M., Moy,M., Murphy,B., Murphy,L.,
Muzny,D.M., Nelson,D.L., Nelson,S.M., Moy,M., Murphy,B., Nixon,K.,
Nusker,D.R., Pacle,J.M., Palazolo,M., Pltman,G.S., Pan,S.,
Pollard,J., Puri,V., Reese,M.G., Reinert,K., Remington,K.,
Saunders,R.D., Scheeler,F., Shen,H., Shue,B.C., Siden-Kiamos,I.,
Simpson,M., Skupski,M.P., Smith,T., Spier,E., Spradling,A.C.,
Stapleton,M., Strong,R., Sun,E., Svirskaas,R., Tector,C., Turner,R.,
Venter,E., Wang,A.H., Wang,X., Wang,Z.Y., Wassarman,D.A.,
Weinstock,G.M., Weissenbach,J., Williams,S.M., Woodage,T.,
Worley,K.C., Wu,D., Yang,S., Yao,Q.A., Ye,Q., Yeh,R.F.,
Zaveri,J.S., Zhan,M., Zhang,G., Zhao,Q., Zheng,L., Zheng,X.H.,
Zhong,F.N., Zhong,W., Zhou,X., Zhu,S., Zhu,X., Smith,H.O.,
Gibbs,R.A., Myers,E.W., Rubin,G.M. and Venter,J.C.

TITLE	The genome sequence of Drosophila melanogaster
JOURNAL	Science 287 (5461), 2185-2195 (2000)
MEDLINE	20196006
REFERENCE	2 (bases 1 to 301639)
AUTHORS	Adams,M.D., Celniker,S.E., Gibbs,R.A., Rubin,G.M. and Venter,C.J.
TITLE	Direct Submission
JOURNAL	Submitted (21-MAR-2000) Celera Genomics, 45 West Gude Drive, Rockville, MD, USA
COMMENT	On Oct 9, 2000 this sequence version replaced gi:7296218.
FEATURES	Location/Qualifiers
source	1. .301639
	/organism="Drosophila melanogaster"
	/db_xref="taxon:7227"
	/chromosome="2L"
	join(2454. .3041,3154. .3514,3593. .4416)
	/gene="CG11023"
	/product="CT9573"
	/db_xref="FLYBASE:FBan0011023"
	/db_xref="FLYBASE:FBgn0031208"
	<2454. .>4416
mRNA	/gene="CG11023"
	/map="21A3-21A4"
	/db_xref="FLYBASE:FBan0011023"
	/db_xref="FLYBASE:FBgn0031208"
	join(2605. .3041,3154. .3514,3593. .4201)
CDS	/gene="CG11023"
	/note="CG11023 gene product"
	/codon_start=1
	/db_xref="FLYBASE:FBan0011023"
	/db_xref="FLYBASE:FBgn0031208"
	/protein_id="AAF51571.1"
	/db_xref="GI:7296280"
	/translation="MGERDQPSSEKISIFNPYYTQHVRNEAPYIPTFDLLSDEESSQVAMAGPSFRPLTYSDAVRLSONGFANSRVSGHSYTVRRPPALVDRSILSOEMERNDQEQYIYLIRTAQNSVSGSHYAEPTDENSEYKVSETNKRLLTRTESLHRRFASCVNLNDFAEQFKARADCEESKHLRLAEORLEFSNESAIKNIDELRAYERKVENIFQSCIAHKPIFVLGPLDKPNVKVTKLIPLTEHHDRNEITQDDKSTETIIFKFNLIHTEDICTEINGKWLNDENVINFYMSLTERSEKRSGLVPATYAINTFEVPRLLQAGHAGIKRWTRKVDLFESKDIIPVPHCVHNCVMCAIHLRNKTIPTYDSKGGKPNRPVLDALFKYLBESIEFKPKKQFDTSFVIESVQNIPIRQLDSDCGIFSCMFAEYITCDVPITFTQS EMLYFRKKMALEIVDDEL"
mRNA	complement(join(4761. .6269,6335. .6443,6704. .7146,7211. .7853,8445. .8550,8608. .9799,9858. .10636,11978. .12137,13186. .13494))
	/gene="1(2)g1"
	/product="CT43225"
	/db_xref="FLYBASE:FBan0002671"
	/db_xref="FLYBASE:FBgn0002121"
	complement(join(4761. .6269,6335. .6443,6704. .7146,7211. .7853,8445. .8550,8608. .9799,9858. .10636,11978. .12137,12951. .13093,13186. .13475))
mRNA	/gene="1(2)g1"
	/product="CT43227"
	/db_xref="FLYBASE:FBan0002671"
	/db_xref="FLYBASE:FBgn0002121"
	complement(<4761. .>16291)
gene	/gene="1(2)g1"
	/note="CG2671"
	/map="21A3-21A4"
	/db_xref="FLYBASE:FBan0002671"
	/db_xref="FLYBASE:FBgn0002121"
	complement(join(4767. .6269,6335. .6443,6704. .7146,7211. .7853,8445. .8550,8608. .9799,9858. .10636,11978. .12137,15990. .16291))
mRNA	/gene="1(2)g1"
	/note="Nucleotide sequence of the Celera sequence differs from the published sequence for this transcript."
	/product="CT9059"
	/db_xref="FLYBASE:FBan0002671"
	/db_xref="FLYBASE:FBgn0002121"
	complement(join(6140. .6269,6335. .6443,6704. .7146,7211. .7853,8445. .8550,8608. .9799,9858. .10636,11978. .12061))
CDS	

/gene="1(2)g1"
/note="1(2)g1 gene product [alt 3]"
/codon_start=1
/db_xref="FLYBASE:FBan0002671"
/db_xref="FLYBASE:FBgn0002121"
/protein_id="AAG22255.1"
/db_xref="GI:10727417"
/translation="MLKFIKGGQGPSADRHRLQKDLFAVYRKTQAQHGFPKPSALAYD PVLKMAIGTQTGALKVFGQPVLEYGQHTLLNNSASELNVQLLEWVYGTGRILSLTA ANOLIMEPVGATLLPIKTLPRFDGKLKKVSSLCSSLDLIMIGTEGNIYQLDLHTE TIKPEVIYHDVVLLEQVPAYKLNPGAIIESIRQLPNSPSKLLVAYNRGLCVLWDFESAS VQRAYIAPHGQSVGLTVNFESEFTWYHADGSYATWSIDNPEPSSNVNVPYGPDPCKSINRLYKGRKRSNDVIVFSGMPRSAYGDHNCVSHASDGHKVCULDFTSKVIDFEFVT FENNRDVAEVLVLLHEELCAVDLPNICAIAKAPYLHSHASAVTCNYLASEVQSV YESILRAGDEQDIDYSNISWPIITGGTLPDNLEESVEEDATKLYEILLTGHEDSYKFW DCTGVLRKPIYNEKTSSIFGSESDRDDAADMASAEQVDEGEPPFRKSGLFDPYSDDP RLAVKKIAFCPKTGQLIVGGTAGQIVIADEFIDLPKVSILKYISMNLVSDRDGFVWKGH DQLNVRSNLIDGEAIPPTTERGVNISGLVQLVLPASITCMALASWGLVSGGTAGLIVL FDFKNFVPEVFERCTLNPNDLTGAGEQLSRKSEFKSLRESFRKLKRGSTRTNQSNQV PPTLEARVERQIEARCADDGLGSMVRCLLFAKTYVTNVNITSPTLWSATNASTVSF LLHLPPAQTAATAVPSASGNAPRPHMRIRISAQLAKEIQLKHRAPVVGISIFDQAGSPV DQNLNAGENGSPHRVLIASEEQRYVESLPOLKPINKYKLTANEGARIRIRIHGFSR ISEPTELQSMHGCSPTKSTRSHGDGEADPNISGSLAVSRGDVYNETALICLTNMGDIMV LSVPELKRQLNAAAVRRREDINGVSSLCFTNSGEALYMSSELSQRIALATSRVQPTG VVPEPLENEESVLEENDAENNKETYACDEVVNTYEIKNPSGISICTRPAEENVRNS VOQVNGVNISNSPNOANETISSIGDITVDSVRDHLNMTTTLCSINTEETIGRSLV STQTNKASTVTVMSEIPNINISNLEDLESKRNTTETSTSSVVIKSIITNISHEKTNGD NKIGTPKTAPEESQF"
complement(join(6140. .6269,6335. .6443,6704. .7146,7211. .7853,8445. .8550,8608. .9799,9858. .10636,11978. .12061))
/gene="1(2)g1"
/note="1(2)g1 gene product [alt 2]"
/codon_start=1
/db_xref="FLYBASE:FBan0002671"
/db_xref="FLYBASE:FBgn0002121"
/protein_id="AAG22256.1"
/db_xref="GI:10727418"
/translation="MLKFIKGGQGPSADRHRLQKDLFAVYRKTQAQHGFPKPSALAYD PVLKMAIGTQTGALKVFGQPVLEYGQHTLLNNSASELNVQLLEWVYGTGRILSLTA

441	levalGlglyArgArgSerleuGlu.....ArgThrThrSer	452
91382	TTACACGACGCTGCTTCGCTTGAAACGGAAGAAAAACACCGATCCTGTGTC	91431
453	GlyAspAspAlaCysAsnLeuThr...SerPheArgProAlaThrLeuTh	468
91432	CCAGATGATTTTGCAAAACGTTGTAGAAAACCTTCGTCCTATCAGTATTAAC	91481
468	rValThrAsnPhenPheGlySgIngluGlyAspArgLeuSerAspGluAspL	485
91482	AGTACCAAGCTTCTTTAAACAAGAGCGAGATAAATGAAGACGAGAGATT	91531
485	eutYrLysPheLeuAlaAspMetArgArgProSerSerValLeuArgArg	501
91532	TGTACAAATTTTACCAGAGTTAAACGACCGAGCTCAGTTATGAAGAAC	91581
502	LeuArgProIleThrAlaGlnLeuLysIleAspIleSerProAlaProGl	518
91582	TACAAATGCATACCTGCTGTTCTATTAATTAGAATATACCTTGCGTCGA	91631
518	uAsnProHisTyrCysLeuThrProGluLeuLeuGlnValLysLeuTyrP	535
91632	GGAAGCCAAGATGCATTAAACGCCAGAGTTAGCACCTATTAACCCCTCAAA	91681
535	roAspSerArgValArgProThrArgGluIleLeuGluPheProAlaArg	551
91682	GTCGAGATTAATGTGCGCCAGTTAAGAAATCTTAGAATTTCCGCAATCG	91731
552	AspValTyrValProAsnThrThrTyrArgAsnLeuLeuTyrIleTyrPr	568
91732	GCGATTTACAATCCACATATATAGCTATCGCAATCTCTATTTGTGTCACC	91781
568	oGlnSerLeuAsnPhenAlaAsnArgGlnGlySerAlaArgAsnIleThrV	585
91782	AAAGGAGCTTAATTTTCTTCGCGTGCAGGTTCTGCTCGAAACATTGCCG	91831
585	AllySValGlnPheMetTyrGlyGluAspProSerAsnAlaMetProVal	601
91832	TTCCGCTTCAACTAATGCGCTGGAGAAACGCCAAAGATGCGGTCAATGCC	91881
602	IlePheGlyLysSerSerCysSerGluPheSerLysGluAlaTyrThrAl	618
91882	ATCTACGGCAAGTCGCTGTGTCCCAAAATTTTCTACTGAAGCATTTACAGC	91931
618	aValValTyrHisAsnArgSerProAspPheHisGluGluIleLysValL	635
91932	TGTCAATTATCATACAAATGCCCATCATTTCTACGATGAATTAATAAATTG	91981
635	ysLeuProAlaThrLeuThAspHisHisLysLeuPheThrPheTyr	651
91982	CGCTGCCTGCATCAATAAAGCAGCATCACCATTTATTGTACCACTTTTAC	92031
652	HisValSerCysGlnGlnLysGln.....AsnThrProLeuGluTh	665
92032	CATGTTTCATGTCAAAAGAAACACACAGACTTGACGCCGTCAGTAGAGAC	92081
665	rProValGlyTyrThrTrpIleProMetLeuGlnAsnGlyArgLeuLysT	682
92082	TCCTATCGGCTACACATGTTGCCCTTACTGGAAGATGAATAACTTAAGT	92131
682	hrgLysGlnPheCysLeuProValSerLeuGluLysProProGlnAlaTyr	698
92132	TTGGGGAATTTAATCTTCCTGTATGTAGTAATCACCCGCGAATAATTAC	92181
699	SerValLeuSerProGluValProLeuProGlyMetLysTrpValAspAs	715
92182	TCATTTTATACCAACCGAATGTTCACTGCGCTGGAATTAATGCTGGACAA	92231
715	nHisLysGlyValAlaPheAsnValGluValAlaValAlaSerSerIleHisT	732
92232	TCACAGAGCCGTGTTTCTATTATGTAGAAACAGTACCGCAATTCATA	92281

732	hrgInAspProTyrLeuAspLysPhePheAlaLeuValAsnAlaLeuAsp	748
92282	CTTTGGATTTCTTCTCGATCGATTCTTCTTAATTGGCAATATCTAGAC	92331
749	GlulHisLeuPheProValArgIleGlyAspMetArgIleMetGluAsnAs	765
:::		
92332	ACTCGAATATATACCTTCCCATATTTGGC.....GAGAATAA	92366
765	nleuGluAsnGluLeuLysSerSerIleSerAlaLeuAsnSerSerGlnL	782
: :	: :	: :
92367	TATAGAAACAGAAATTAAAGAAATGCTTGTGATATCGAATATGCAATC	92416
782	euGIuProValValArgPheLeuHisLeuLeuLeuAspLysLeuIleLeu	798
92417	GTCAGCCATTAGTAGAGGCATCTACCATTTAGTCTTGGATTAAGCTTATCGAA	92466
799	LeuValIleArgProProValIleAlaGlyGlnIleValAsnLeuGlyI	815
: :	: :	: :
92467	TTGTTAGTTGTTCACCCACAAAGTTGGTGACAAAGCAATGTCTTGGAATC	92516
815	nAlaSerPheGluAlaMetAlaSerIleIleAsnArgLeuHisLysAsnL	832
92517	TACGGTTTTCGAAGTCTTGTGTTGGTATCATCTTGTGATCAATACTTA	92566
832	euGIuGlyAsnHisAspGlnHisGlyArgAsnSerLeuLeuAlaSerTyr	848
: :	: :	: :
92567	ACGACGATCATATATGACCAATACGGTTCGCAAAAGTTGCTATCAACATAT	92616
849	IleHisTyrValPheArgLeuProAsnThrTyrProAsnSerSerSerPr	865
: :	: :	: :
92617	GTGCACCTTTCATGTAAATTTCCGCACCCCTTCCA.....	92652
865	oGIuProGlyGlyLeuGlyGlySerValHisTyrAlaThrMetAlaArgS	882
92653AGCAACACGCGACTTAACATGCAGTCGCA	92680
882	erAlaValArgProAlaSerLeuAsnLeuAsnArgSerArgSerLeuSer	898
: :	: :	: :
92681	GCACAACAGAA.....GACTTGCAATTAAAGCAATGTTACACCATATAT	92724
899	AsnSerAsnProAspIleSerGlyThrProThrSerProAspAspGluVa	915
92725	GATAAT.....	92730
915	lArgSerIleIleGlySerLysGlyLeuAspArgSerAsnSerTrpValA	932
: :	: :	: :
92731	...GTTTTAGCAAGTGGCCGAAAGCTTGATCGCAAAAGTTAA.....A	92769
932	snThrGlyGlyProLysAlaAlaProTrpGlySerAsnProSerProSer	948
92770	ATACATATTTTGTAAAGCT.....	92789
949	AlaGluSerThrGlnAlaMetAspArgSerCysAsnArgMetSerSerHi	965
92790TTAAATATAACCTTTTCT	92809
965	stHrgIuThrSer...SerPheLeuGlnThrLeuThrGlyArgLeuProT	981
: :	: :	: :
92810	TTCAAGAGCTGTCCATAGATGTCTTCACTCGATGGCTGCACGGGATGTCC	92859
981	hrLysLysLeuPheHisGlyGluLeuAlaLeuGlnTrpValCysSer	997
: :	: :	: :
92860	AAGTGGCGTGTTCACGAAGAATTAGCCCTTGCAATTGGGTGTGCAAGT	92909
998	GlySerValArgGluSerAlaLeuGlnGlnAlaTrpPhePheGluLe	1014
: :	: :	: :
92910	GGAAAGCAGCTGACTTAGCCCATGTCAAAATTCGTGTTTATTTGAGCT	92959
1014	uMetValLysSerMetValHisHisLeuTyrPheAsnAspLysLeuGluA	1031
: :	: :	: :
92960	TATGTCAAAATCTATGATTGAACATTTGCATTGTCAAAATACTTTAAATG	93009
1031	lapProArgLysSerArgPheProGluArgPheMetAspPileAlaAla	1047

93010 GTCCCCGCAACATCGATTTCACATCAGTTTAATGATGATCTATCCACC 93059
1048 LeuValSerThrIleAlaSeraspIleValSerArgPheGlnLysAspTh 1064
93060 CTGGTGCACCTTAGTTACACAAAGTGGTCGGCTACCACACGCAACGACC 93109
1064 rGlumetValGluArgLeuAsnThrSerLeuAlaPhePheLeuAsnAspL 1081
93110 TAAATTAGCGCAGTCATTAATGCAAGTTTAGCCTTTTATATTGACA 93159
1081 euLeuSerValMetAspArgGlyPheValPheSerLeuIleLysSerCys 1097
93160 TACTAAGCATCATGGATCGGGATTGTTTCGGCCTCATCAAAACGTAC 93209
1098 TyrLysGlnValSerSerLysLeuTyrSerLeuProAsnProSerValLe 1114
93210 ACAAGGTTCTGATTTCTAAAAATGCTTCAATACCGAT.....CT 93250
1114 uValSerLeuArgLeuAspPheLeuArgIleIleCysSerHisGluHist 1131
93251 AATGAACATAAAATAGACTTTTGAAGATTGTGTAGTCATGAGCATT 93300
1131 yrValThrLeuAsnLeuPro.....CysSerLeuLeuThrPro 1143
93301 TTGTAGCCTTAAACTTGCCGTTTGCTACTTCATACACAATGTGAACAGCG 93350
1144 ProAlaSerProSerProSerValSerSerAlaThrSerGlnSerSergl 1160
93351 CCATGACGCCCCACGCGCAAGCACACGTCGACACAGCAACGCAACTTCTTG 93400
1160 ypHeSerThrAsnVal..... 1165
93401 CGTAAGTAAAAACAATACACTTAAATTTTACTTATTTGGCTTCTT 93450
1166GlnAspGln..LysIleAlaAsnMetPheGluLeuSerValProp 1180
93451 TTAAAGGATCAGTACAGAGAGCCCTTCACGCCGACCTTAGTCAAGAAT 93500
1180 heArgGlnGlnHisTyrLeuAlaGlyLeuValLeuThrGluLeuAlaVal 1196
93501 TCCTCCAGCACATTTTCTAGTCGACTTGTCTGAGCGACTTAGCAGCA 93550
1197 IleLeuAsp...ProAspAlaGluGlyLeuPheGlyLeuHisLysLysVa 1212
93551 GTAATGAGGTGCCAAACCCCAA.....TTACATGGAAGAAAGC 93588
1212 IleAsnMetValHisAsnLeuLeuSerSerHisAspSerAspProArgT 1229
93589 AATAAGGTGTATTCGCAACCTGATGACATCTCAGATTGTGGACGCGCGGT 93638
1229 yrSerAspProGlnIleLysAlaArgValAlaMetLeuTyrLeuProLeu 1245
93639 ATAGCGAAACTGATGCACGTGCAAGGTGGCTGCTTATATATACCGCTA 93688
1246 IleGlyIleIleMetGluThrValProGlnLeuTyrAspPhe..... 1259
93689 CTTTCTATAGATGATGATTCGATTCCTCAGTTGCATCAGCACGCGCTTGA 93738
1260 .ThrGluThrHisAsnGlnArgGlyArgProIleCysIleAlaThrAspa 1276
93739 GCAGGATCGTTCGACGCAAAATCGGTCAA.....CTTGAAG 93773
1276 spTyrGluSerGlu.....SergLysSerMetIleSerGlnThr 1288
93774 ACTATCAAGGGCCACATCAAACTATTTACAGCATCAACCATAAATCCGGAA 93823
1289 ValAlaMetAlaIleAlaGlyThrSerValProGlnLeuThrArgProGl 1305
93824 GTGGCATTTGCGATATCGGGA.....AGTCGCCCTTA 93855
1305 ySerPheLeuLeuThrSerThrSergLysArgGlnHisThrThrPheSera 1322

93856 CTCCTACTTG.....AATGAACAAGTTAAAAACAATTGCCCCCTCAGCT 93899
1322 IaGluSerSerArgSerLeuLeuIleCysLeuLeuTrpValLeuLysAsn 1338
93900 CGGAGAAATACCCGCATTTATTGTGTGTCTGTGTGGTGAATAAAAC 93949
1339 AlaAspGluThrValLeuGlnLysTrpPheThrAspLeuSerValLeuGl 1355
93950 TTGGAGCGCAACGTTTATATCGCTGCGTTATGATTTAAGTCCGCATAG 93999
1355 nLeuAsnArgLeuLeuAspLeuLeuTyrLeuCysValSerCysPheGluT 1372
94000 AGTGCACCAATGCTTCAAGGTGTTAACGTTTGCTTGAAACTTTCGAGT 94049
1372 yrLysGlyLysLys.....ValPheGluArgMetAsnSerLeuThr 1385
94050 ACACAGGACAAAAAACGTGGCCACCCTTAAACGACGAAATACACAAAGT 94099
1386 PheLysLys.....SerLysAspMetArgAlaLysLeuGluGluAlaI 1400
94100 TTTAGGAAAACGTGATCTACTGATGTTAAGGAAAAAGCTGGAAGAAATGTAT 94149
1400 eleuGlySerIleGlyAlaArgGlnGluMetValArgArgSerArgGlyG 1417
94150 TAGAGGCACAAATTCAGCACGTTATGATTTAATAATCGTCGAAAAAGATC 94199
1417 InLeuGluArgSerProSerGlySerAlaPheGlySergLysGlnLysLeu 1433
94200 GT.....AATTCACAGAAAAAATTC 94219
1434 ArgTrpArgLysAspMetThrHisTrpArg...GlnAsnThrGluLysLe 1449
94220 CGATGGAGGAAGACACAGATGCCATATCCGCTCTCAATATGCGGATGCTGT 94269
1449 uAspLysSerArgAlaGluIleGluHisGluAlaLeuIleAspGlyAsnL 1466
94270 GGGAAAAAGTGAGCATGAACCTGAATTAAGTCATTTTCATCGAAGGCTCTT 94319
1466 euAlaThrGluAlaAsnLeuIleIleLeuAspThrLeuGluIleValVal 1482
94320 TGGCAACTGAAGTTCGACTGTGTACTCTCGATACCCCTTGAAATTAATGTT 94369
1483 GlnThrValSerValThrGluSerLysGluSerIleLeuGlyGlyValLe 1499
94370 CAT.....GCTGCTGCCAACCTTATACAAATCTTCTTGGAACGTGCT 94413
1499 uLysValLeuLeuHisSerMetAlaCysAsnGlnSerAlaValTyrLeuG 1516
94414 AAAGGTGCTTCTTCATTTCTTGTTCGCCGAATCAATCTGTTTGGCCCTGC 94463
1516 InHisCysPheAlaThrGlnArgAlaLeuValSerLysPheProGluLeu 1532
94464 AAAACTTGTGCTGCCCAACGCGCGTTAATCTTTAAATTTCCAAACTTA 94513
1533 LeuPheGluGluGluThrGluGlnCysAlaAspLeuCysLeuArgLeuLe 1549
94514 TTATTGTGATGACGAGACCGATATATGTGCGGATTGTGTCTAATTTTACT 94563
1549 uArgHisCysSerSerSerIleGlyThrIleArgSerHisProSerAlas 1566
94564 GAAACATTTGTGGTCCCTATTTGCCAGGAATACGATCACAGCAGCTGCAT 94613
1566 erLeuTyrLeuLeuMetArgGlnAsnPheGluIleGlyAsn..... 1579
94614 CACTATATTACTAATGAGACAAAAATTTGAAATTTGGAATGTAAAGTTT 94663
1579 1579
94664 ATTTTATTGTTTAGGTTTATTAACTAATTAATGATGTCACCTTATTCGTAT 94713
1580 .Asn.PheAlaArgValLysMetGlnValPrometSerLeuSerSerLeu 1595
94714 AGAACTTTGCCGAGTTAAGATGCAAGTGACGATGTCTTTAAGCTCCCTA 94763

1596 ValGlyThrSerGlnAsnPheAsnGluGluPheLeuArgSerLeu 1612
|||||
94764 GTTGGACACAGTTCGGTCTTTAGTGAGCAATCCCTTACGCCGTGCGCTAA 94813
1612 sThrIleLeuThrTyrAlaGluGluAspLeuGluLeuArgGluThrThrP 1629
|||||
94814 AACCGTCCGTGTTATGCTGAATCCGACTCCGACTCGAGATGATACTTCTT 94863
1629 hepProAspGlnValGlnAspLeuValPheAsnLeuHisMetIleLeuSer 1645
|||||
94864 TTCCTGAACAAGTACAGATTGCTTTTAACTGCATATGATCCTGTCA 94913
1646 AspThrValLysMetLysGluHisGlnGluAspProGluMetLeuIleAs 1662
|||||
94914 GATACTGTTAAAATGAAGAGTATCAGGAAGACCCAGAAATGTTGCTTGA 94963
1662 pleuMetTyrArgIleAlaLysGlyTyrGlnThrSerProGlu...ArgL 1678
|||||
94964 CCTCATGAATCGTATGCGCAAGGATACCAAAATAACCTGATCTACGAC 95013
1678 euThrTrpLeuGlnAsnMetAlaGlyLysHisSerGluArgSerAsnHis 1694
|||||
95014 TGACTTGGTTGGAATAATGCGCTAAAAAACACCCGCGAGCAAAATCAC 95063
1695 AlaGluAlaAlaGlnCysLeuValHisSerAlaAlaLeuValAlaGluTyr 1711
:::
95064 ACGGAAGCAGCCATGTGTATGTACATGCTGCTCTTTAGTTTCTGAATA 95113
1711 rLeuSerMetLeuGlnAspArgLysTyrLeuProValGlyCysValThrP 1728
|||||
95114 TCTTAGCATGTTGAGTCACAAAACATTTGCTGTGGAGCTGTAAGTT 95163
1728 heGlnAsnIleSerSerAsnValLeuGluGluSerAlaValSerAspAsp 1744
|||||
95164 TTCAACGAATTTCTCCCAACACACTAATGAGTCGCCGTATCGGATGAT 95213
1745 ValValSerProAspGlnGluGlyIleCysSerGlyLysTyrPheThrG 1761
|||||
95214 GTGCTAAGTCCCGCGAGAGATGATCTGCCCTAGGAAATCATTTCACTGA 95263
1761 uSerGlyLeuValGlyLeuLeuGlnAlaAlaAlaSerPheSerMetAla 1778
:::
95264 AACTGGGTTGAAGGCCCTTGCTGGAAGAGCCCTCCAATTTCTTCAAGTTG 95313
1778 laGlyMetTyrGluAlaValAlaAsnGluValTyrLysValLeuIleProIle 1794
|||||
95314 CTGGCATGTATGAAGCAATGAACGAAGTGTACAAAAATTTCTAATACCATTA 95363
1795 HisGluAlaAsnArgAspAlaLysLysLeuSerThrIleHisGlyLysLe 1811
|||||
95364 TCGGAGGCTAACAGAGATTTTCAAAAGCTAAGCAAAAGTTCATGGCAATT 95413
1811 uGlnGluAlaPheSer..... 1816
|||||
95414 GCAGGAGGCATTTAATCGAATATCCCAACTACAGGTAAACAATATTGTGTA 95463
1817LysIleValHisGlnSerThrGlyTyrGluArgMe 1828
|||||
95464 AATTTTACCACGGAATAATATATACATATTATAAACAGGTAAGAGAGT 95513
1828 tPheGlyThrTyrPheArgValGlyPheTyrGlyThrLysPheGlyAspL 1845
:::
95514 TTTTGGACATACTTTCGTGTGCGCTTCTATGCGGAAAAATTTGGGGAAT 95563
1845 euAspGluGlnGluPheValTyrLysGluProAlaIleThrLysLeuAla 1861
:::
95564 TGGATCAGCAGGAATTCATTATAAAGAGCCAACATTGACGAAGTTGCC 95613
1862 GluIleSerHisArgLeu.Glu..... 1868
|||||
95614 GAAATATTATTAGTCGCTTCAGGTATATATTGCAAAATTGGAATAATAGAA 95663

1869GlyPheTyrGlyGluArgPhe 1875
95664 CTAATCAATTTTGTTCACATACGTTAGAACCTTTTACACTGAACGATTC 95713
1876 GlyGluAspValValGluValIleLysAspSerAsnProValAspLysCys 1892
|||
95714 GGACCGGACTCTGTGCATATCATTTAAAGATTCCAAATACCGGTGATATTAA 95763
1892 sLysLeuAspProAsnLysAlaTyrIleGlnIleThrTyrValGluProT 1909
:::
95764 TAGCTTGATCCCGATTAAGGCTTACATTCAAAATTACTTATGTTGAACCT 95813
1909 yrPheAspThrTyrGluMetLysAspArgIleThrTyrPheAspLysAsn 1925
|||||
95814 ACTTGAACATATGAATGCGTCATCGTGAGACATACTTTGAGCGGAAT 95863
1926 TyrAsnLeu..... 1928
:::
95864 TTCAATATTAAGTATGATATGAATTAAACAGATAATTTAATCGAAATTTA 95913
1929Arg.ArgPheMetTyrCys 1934
95914 AAATATTAATTTTACATTTTGGTTTATGACAGAACGTTTATATATGCCC 95963
1935 ThrProPheThrLeuAspGlyArgAlaHisGlyGluLeuHisGluGlnP 1951
|||||
95964 ACGCCTTTTACTAAAAATGGAAGCCACATGGCGCACTTAATGAGCAGTGC 96013
1951 eLysArgLysThrIleLeuThrThrSerHisAlaPheProTyrIleLysT 1968
|||||
96014 TAAACGGAACACGATTTTGACACAGGCGGAATCATTTCATACGTAAAAA 96063
1968 hrArgValAsnValThrHisLysGluGluIleIleLeuThrProIleGlu 1984
|||||
96064 CTGCGATTATGTTAATTAGCGCAGACAGCAATAGTCTGGAACCGATTGAA 96113
1985 ValAlaIleGlu..... 1988
|||||
96114 GTGGCAATTGAAGTAAGTTAAGTTGTAATAATCCCAAAAGTTAATAAAATT 96163
1989AspMetGlnLysLysThrGlnGluLeuAlaPheA 2000
96164 AAACGTTATTGCCCTAGAGATTTCAAAAAAACAATTTGGAATTTGCCAGCTG 96213
2000 laThrHisGlnAspProAlaAspProLysMetLeuGlnMetValLeuGln 2016
|||||
96214 CCACGAATCAAGAGCCAGCAGACCCCAAAATATTGCCAAATGCTACTGCAG 96263
2017 GlySerValGlyThrThrValAlaAsnGlnGlyProLeuGluValAlaGlnVa 2033
|||
96264 GGATGTATTGGAACCACTGTTAACCAAGACCGATGGAATGGCAGATGT 96313
2033 lPheLeuSerGluIleProSerAspProLysLeu...PheArgHisHisA 2049
|||||
96314 GTTCCTTTCCAAATTATCGAGCGGAACAACCTGTACGCAAAAAACACCAAA 96363
2049 snLysLeuArgLeuCysPheLysAspPheThrLysArgCysGluAspAla 2065
|||||
96364 ACAAACTTCGGTTGTGCTTCGCGAGTTTCAAAAGCGTTGTGCTGATGCT 96413
2066 LeuArgLysAsnLysSerLeuIleGlyProValGlnLysGluTyrGlnAr 2082
:::
96414 TTGAAAAAGAATCGCAATCTAATACTTTACAGATCAAAAAGATTACCAACG 96463
2082 gGluLeu 2084
|||||
96464 AGAACTG 96470

seq_name: gb_pr:AK055905
seq_documentation_block: 3327 bp mRNA linear PRI 31-OCT-2001
LOCUS AK055905
DEFINITION Homo sapiens cDNA FLJ31343 f1s, clone MESAN1000101, weakly similar

ACCESSION AK055905 to Rat trg gene product.
VERSION AK055905.1 GI:16550749
KEYWORDS oligo capping, fls (full insert sequence).
SOURCE Homo sapiens normal mesangial cells (NHMC56046-2) cDNA to mRNA, clone_lib:MESAN1 clone:MESAN1000101.
ORGANISM Homo sapiens
REFERENCE 1 (sites)
AUTHORS Tashiro,H., Yamazaki,M., Watanabe,K., Kumagai,A., Itakura,S., Fukuzumi,Y., Fujimori,Y., Komiyama,M., Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H., Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K., Wagaatsuma,M., Murakawa,K., Kanehori,K., Takahashi-Fujii,A., Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y., Sugano,S., Nagahari,K., Masuho,Y., Nagai,K. and Isogai,T.
TITLE NEDO human cDNA sequencing project
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 3327)
AUTHORS Isogai,T., Otsuki,T. and Sugiyama,T.
TITLE Direct Submission
JOURNAL Submitted (24-OCT-2001) Takao Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:genomics@hri.co.jp, Tel:81-438-52-3951, Fax:81-438-52-3952)
COMMENT NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: RAB and HRI.
FEATURES
source location/Qualifiers
1..3327
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MESAN1000101"
/cell_type="normal mesangial cells (NHMC56046-2)"
/clone_lib="MESAN1"
/note="cloning vector: pME18SFL3-primary culture, normal mesangial cells"
54..2279
/note="unnamed protein product"
/codon_start=1
/protein_id="BAB71042.1"
/db_xref="GI:16550750"
/translation="MELSVPERQOHYLAGIVLELAVILDPDAEGLEGHKKVINMVHNLSSHSDPRYSDPQIKARVAMLYPLIGIMETVPOLYDFTETHNQGRPICIATDDYSESGSMISQTYAMAIAIGTSVPQLTRPGSFLLTSTSGRQHTTFSAESSRLLICL LMWLKNADETVLQKWFDTLSVLQNLRLDLVLCVSCFEYKGKVFERNMSLTFKSK DMRAKLEEAIIIGSIGARQEMVRRSRGOLERSPSGSAFSGOENLRWRKDMTHWRQNTK LDKSRAEIEHEALIDGNLATEANLIIDTLEIVYQTVSVTESKESILGVEKVLHSM ACNOSAVYLQHCFAQRALVSKPELLFEETRECCADCLRLRHCSSTIGTIRSHAS ASLYILMRONFEIGNNFARVKMQVTMSLSISVTSQNFNEEFLRRSLKTLITYAEDL ELRETFPDQVQDLVFNLMILSDTVKMKHEQEDPEMLIDLMYRIAKGYQTSPLRLT WLONMAGKHSERSNHAEAAOCIVHSAALVAEYLSMLEDRKYLPGCVTFONISSNVLE ESAVSDDVSPDEEGICSGKYFTESGLVGLLEQAAASFSMAGMEAVANEYKVLPIH EANRDAKLISTIHGKLGAEAFSKIVHODGRMFGEYFRVGFYGTKEFGDLQGEFVYKEP AITKLAIEISHRLEGFYGRFEGDEVVEVIKDSNPYDKCKLDPNTHIFRLPMWHTLTLM R"
BASE COUNT 1046 a 640 c 711 g 930 t
ORIGIN
alignment_scores:
Quality: 4674.50 Length: 937
Ratio: 5.043 Gaps: 3
Percent Similarity: 98.933 Percent Identity: 98.613
alignment_block:
US-09-737-246-2 x AK055905 ..

Align seg 1/1 to: AK055905 from: 1 to: 3327
1156 SerGlnSerSerGlyPheSerThrAsnValGlnAspGlnLysIleAlaAs 1172
|||||
3 TCTCAGAGTCTCGATTTTCTACGAAATGTACAAGACCAAGAAAGATTGCAAA 52
1172 nMetPheGluLeuSerValProPheArgGlnGlnHisTyrLeuAlaGlyL 1189
|||||
53 TATGTTGAATTAATCCGTGCCTTCCGCCAACAGCATTAATTGGCAGAGAC 102
1189 euValLeuThrgluLeuAlaValIleLeuAspProAspAlaGluGlyLeu 1205
|||||
103 TTGTGTTAACAGAGCTGGCTGTCAATTTAGAACCCGTGATGCTGAAGACTG 152
1206 PheGlyLeuHisLysLysValIleAsnMetValHisAsnLeuLeuSerSe 1222
|||||
153 TTTGGATTGCATAGAAAGTCATCAATATGTGTACACAAATTTACTCTCCAG 202
1222 rHisAspSerAspProArgTyrSerAspProGlnIleLysAlaArgValA 1239
|||||
203 TCACGACTCAGACCCGCGGTACTCTGACCCCTCAGATAAAGGCTCGAGTGC 252
1239 laMetLeuTyrLeuProLeuIleGlyIleIleMetGluThrValProGln 1255
|||||
253 CCATGTGTATCTACCTCTGATTTGGTATATCATGAAGAACTGTACTCTAG 302
1256 LeuTyrAspPheThrgluThrHisAsnGlnArgGlyArgProIleCysIL 1272
|||||
303 CTGTATGATTTTACAGAAACTCACAATCATCAGAGGAAGAACCAATTGTAT 352
1272 eAlaThrAspAspTyrGluSerGluSerGlySerMetIleSerGlnThrV 1289
|||||
353 AGCCACTGATGATTATGAAGTGAGAGCGGAAGTATGATAAGCCAGACCG 402
1289 aAlaMetAlaIleAlaGlyThrSerValProGlnLeuThrArgProGly 1305
|||||
403 TTGCCATGGCAATCGCAGGACATCGGTCCCTCAACTAACAAAGGCTGCG 452
1306 SerPheLeuLeuThrSerThrSerGlyArgGlnHisThrThrPheSerAl 1322
|||||
453 AGTTTCCCTCCTCACGTCACAGAGTGGCAGGCAACACACTACCTTTTCAGC 502
1322 agLusSerSerArgSerLeuLeuIleCysLeuLeuTyrValLeuLysAsnA 1339
|||||
503 AGAATCAAGTCGAAGCCTTTTGATCTGTCTACTTTGGGTTCTCAAAAATG 552
1339 laAspGluThrValLeuGlnLysTrpPheThrAspLeuSerValLeuGln 1355
|||||
553 CAGATGAACAGCTTCTACAGAAAGTGGTTTACAGATCTCTCAGTCTTGCAG 602
1356 LeuAsnArgLeuLeuAspLeuLeuTyrLeuCysValSerCysPheGluTyr 1372
|||||
603 CTAAACCGGCTATTAGATCTGCTTTATCTCTGTGTCTTGCTTTGAGTA 652
1372 rLysGlyLysLysValPheGluArgMetAsnSerLeuThrPheLysLys 1389
|||||
653 TAAAGGGAAGAAAGTGTGTGAACGAATGAATAGCTTGACCTTTAAGAAAT 702
1389 erLysAspMetArgAlaLysLeuGluGluAlaIleLeuGlySerIleGly 1405
|||||
703 CAAAGACATGAGAGCAAGCCTTGAAAGACTAATTCTTGGAGACATAGGT 752
1406 AlaArgGlnGluMetValArgSerArgGlyGlnLeuGluArgSerPr 1422
|||||
753 GCCAGGCAAGAAATGTTAGGGCGGAAGCCGAGACAGCTCGAGAGAAGCCC 802
1422 oSerGlySerAlaPheGlySerGlnGluAsnLeuArgTrpArgLysAspM 1439
|||||
803 ATCTGGAAGTGCCCTTGTGAAGTCAAGAAATTTGAGGTGAGAGAAAGATA 852
1439 etThrHisTrpArgGlnAsnThrGluLysLeuAspLysSerArgAlaGlu 1455
|||||

853 TGACTCACTGGCGTCAAAACACAGAGAAGCTTGACCAATCAAGACGACAG 902
1456 IIEGLUHSGLUALALEUIEASPGLYASNLEUALATHRGUALAASNLE 1472
|||||
903 ATTGAACACGAAGCACTGATTGATGGAACCTGGCTACAGAAGCAAACTT 952
1472 UIIEILEUASPTHREUGLUILEVALVAILGNTHRVASERVALTHRG 1489
|||||
953 AATCATTTTAGATACATTAGAGATTGTTGTTACAGACCGTTTCTGTAACGG 1002
1489 IUSERLYSGUSERILEUGLYGLYVALLEULYSVALLEUENHISSE 1505
|||||
1003 AATCCAAAGAGAGCAATTCTGGTGAGTGTAAAGTGCTACTACACAGC 1052
1506 METALACYSASNGLINSEVALVALTYRLEUGLNHISCYSPHEALATHRG 1522
|||||
1053 ATGGCCTGTAAACCAAGTGCAAGTTATCTACACACTGTTTGGCTACACA 1102
1522 NARGALALEUVALSERLYSPHEPROGLULEUENPHEGLUGLUUUTHRG 1539
|||||
1103 GAGAGCCTTGTTCAAGTTTCCTGAACTCTTATTGGAAGAGAGACAG 1152
1539 LUGLNCYSALAASPLEUCYSLEUARGLEUENARGHISCYSSERSE 1555
|||||
1153 AGCAGTGTGCTGATTTATGCCCTCAGGCTTCTCCGACACTGTAGCAGTAGC 1202
1556 IIEGLYTHRILEARSERHISPROSERALASERLEUTYRLEUENMETAR 1572
|||||
1203 ATCGGTACAATACGGTCCACAGCCAGTGCCTCCCTTACTACTAATGAG 1252
1572 GGLNASNPHEGLUIIEGLYASNASNPHLEAARGVALLYSMETGLNVALP 1589
|||||
1253 GCAAAACTTTGAGATTGGGAATAACTTGGCCAGGTTAAATGCAAGTTAA 1302
1589 ROMETSERLEUSERSERLEUVALGLYTHRSERGLNASNPHEASNGLU 1605
|||||
1303 CAATGTCACTATCCCTCCTGGTGGGCACTCTCAGAATTTTAAATGAAGAA 1352
1606 PHELEUARGARGSERLEULYSTHRIIELEUTHRTYRVALAGLUASPLE 1622
|||||
1353 TTCTTAAGACGTTCTCTAAAGACTATATTGACATATGCTGAAGAAGATCT 1402
1622 UGLULEUARGLUUTHRTHRPHROASPGLNVALGLNASPLEUVALPHEA 1639
|||||
1403 GGAATTGAGGGAACAACATTTCTCGATCAGGTCAGAGATCTGGTTTCA 1452
1639 SNLEUHSMETILEUSERASPTHRVALLYSMETLYSGLUHISGLINGLU 1655
|||||
1453 ATCTCCATATGATTTCTTCTGATACTGTGAAAATGAAGGAACACCAGAG 1502
1656 ASPPROGLUMETLEUIEASPLEUMETTYRARGILEALALYSGLYTYRGL 1672
|||||
1503 GATCCTGAAATGTTGATTGATCTAATGTACAGAATTGCCAAGGGTTACCA 1552
1672 NTHRSERPROGLU...ARGLEUTHRTPLEUGLNASNMETALAGLYLYSH 1688
|||||
1553 GACCTCTCCAGATCTGCGACTGACTGGTTGCAGACATGGCAGGCAAGC 1602
1688 ISSERGLUARGSERASNHISALAGUALAALAGLNCYSLEUVALHISSE 1704
|||||
1603 ACTCAGAACGAACAATCATGCTGAAGCTGCACAGTGTCTAGTCCACTCA 1652
1705 ALAALALEUVALALAGLUITYRLEUSERMETLEUGLUASPARGLYSTYRLE 1721
|||||
1653 GCAGCACTTGTGCTGAATATTGAGCATGCTGGAGGACCGGAATATCT 1702
1721 UPROVALGLYCYSVALTHRPHENASNILESERSEASNVALLEUGLU 1738
|||||
1703 TCTGTGGGATGTGTAACATTTCAGAATATTTCATCTAATGTTTAGAAG 1752
1738 IUSERALAVASERASPARVALVALSERPROASPGLUUGLYILECY 1754
|||||
1753 AATCTGCGGCTCTCAGATGATGGTATCTCCAGATGAGAAGGTATCTGC 1802

1755 SERGLYSTYRPHETHRGUSERGLYLEUVALGLYLEUENGLUGLNAI 1771
|||||
1803 TCTGAAAATACTTTACTGTAGTCAGGACTTGTGGGATTAATGAACAAGC 1852
1771 AALAASERPHESERMETALAGLYMETTYRGLUVALVALASNGLUVALT 1788
|||||
1853 AGCTGCTTCTCTCTPATGGCTGGCATGTATGAGCAGTTAATGAAGTTT 1902
1788 YRYSVALLEUILEPROILEHISGLUALAASNARGASPALALYSLSLEU 1804
|||||
1903 ACAAGTACTTATTCCTATTCATGAAGCTAATCGGGATGCAAGAAGACTA 1952
1805 SERTHRIIEHISGLYLSLEUGLNUALALAPHESERLYSILEVALHISGL 1821
|||||
1953 TCCACAATTTCATGGTAAACTTCAGAAGCATTCACCAAAATTGTTCAATCA 2002
1821 NSERTHRGLYTRPGLUARGMETPHEGLYTHRTYRPHARGVALGLYPHET 1838
|:::|
2003 GGAT....GGTAAGCGGATGTTGGCACCTTAATTTTCGTGTTGTTT 2046
1838 YRGLYTHRLYSPHEGLYASPLEUASPGLUINGLUPHLEVALTYRLYSGLU 1854
|||||
2047 ATGGAACCAAGTTGGGGATTTGGATGAACACAAATTTGTTTACAAGGAG 2096
1855 PROALAIETHRLYSLEUALAGLUILESERHISARGLEUGLUGLYPHET 1871
|||||
2097 CCTGCAATTAACCAACTTGCAGAGATATCTCAGAGATTGGAGGATTTTA 2146
1871 RGLGLUARGPHEGLYGLUASPARVALGLUVALILELYSASPSEASN 1888
|||||
2147 CGGAGAAAGATTGGAGAGGATGTGTTGAAGTAATCAAGACTCTAATC 2196
1888 ROVALASPLYSCYLSLEUASPPROASNLYSALATYRILEGLINLETHR 1904
|||||
2197 CTGTAGACAGTGTAATTAGATCCTAACAC.GCATATATTACAGATTACC 2245
1905 TYRVALGLUPROTYRPHASPTHRTYRGLUMETLYSASPARGILETHRTY 1921
|||||
2246 TATGTGAGCCATACTTTGACACATATGAGATGAAGGACAGAAATCACTTA 2295
1921 RPHEASPLYASNTYRASNLEUARGARGPHEMETTYRCYSTHRPROPHET 1938
|||||
2296 TTTCGACAAAAATTAACAATCTTCTGTCATTCATGTACTGTACACCCCTT 2345
1938 HRLEUASPGLYARGALAHISGLYGLULEUHSGLINGLNPHELYSARGLY 1954
|||||
2346 CTTTAGATGGCCGTCCTCATGGGAACCTTCATGAACAATTCAAAAGGAAG 2395
1955 THRILEUTHRTHRSERHISALAPHEPROTYRIELYSTHRARGVALAS 1971
|||||
2396 ACCATTCTGACACGTCTCATGCTTCTTATATTAATAACAAGGCTCAA 2445
1971 NVALTHRHISLYSGLUUIEILELEUTHRPROILEGLUVALALILEG 1988
|||||
2446 TGTCACTCATTAAGACAAGATCATCTTAACACCAATTGAAGTGTCTATTG 2495
1988 LUASPMETGLNLYSTYRTHRGINGLULEUALAPHEALATHRHISGLNASP 2004
|||||
2496 AGGACATGCAGAAAAAGACACAGAGATTGGCAATTTGCCAACACATCAGGAT 2545
2005 PROLAASPPROLYSMETLEUGLIMETVALLEUGLINGSERYVALGLYTH 2021
|||||
2546 CCCGAGACCCCAAAATGCTTCAGATGTACTCCAGGATCTGTAGGCAC 2595
2021 RTHRVALASNGLINGLYPROLEUGLUVALALAGLNUALPHLEUSERGLUI 2038
|||||
2596 CACAGTGAATCAGGGCCTTTGGAAGTTGCCACAGGTTTTCGTCTGAAA 2645
2038 LEPROSERASPPROLYSLEUPHEARGHISHISASNLYSLEUARGLEUCYS 2054
|||||
2646 TACCTAGTGACCCAAAGCTCTTCAGACATCATATAATAACTGCGACTCTGC 2695

2055 PhelysAspPheThrLysArgCysGluAspAlaLeuArgLysAsnLysSs 2071
|||||
2696 TTTAAAGATTCTACTAAAGGTGTGAAGATGCTTTAAGAAAAAATAAGAG 2745
|||||
2071 rLeuIleGlyProValGlnLysGluTyrGlnArgGluLeu..GlyLysLe 2087
|||||
2746 CTTAATTTGGCCGATCAAAAGAGATATCAAAGGGAAGCTGAGAGAAACT 2795
|||||
2087 userSerPro 2090
|||||
2796 ATCATCGCCT 2805
seq_name: gb_pat:AX172978

seq_documentation_block: 3253 bp DNA linear PAT 03-JUL-2001
LOCUS AX172978
DEFINITION Sequence 105 from Patent WO0142297.
ACCESSION AX172978
VERSION AX172978.1 GI:14597957
KEYWORDS
SOURCE synthetic construct.
ORGANISM artificial sequence.
REFERENCE 1 (bases 1 to 3253)
AUTHORS Lu,P., Garman,J.D. and Candia,A.F.
TITLE Clasp-3 transmembrane protein
JOURNAL Patent: WO 0142297-A 105 14-JUN-2001;
Arbor Vita Corporation (US)
FEATURES
source 1..3253
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="polynucleotide fragment"
BASE COUNT 1009 a 633 c 705 g 906 t
ORIGIN

alignment_scores: Quality: 4626.50 Length: 904
Ratio: 5.123 Gaps: 1
Percent Similarity: 99.889 Percent Identity: 99.558

alignment_block:
US-09-737-246-2 x AX172978 ..
Align seg 1/1 to: AX172978 from: 1 to: 3253

1188 GLEuValLeuThrGluLeuAlaValIleLeuAspProAspAlaGlu 1204
|||||
4 GGACTTGTGTTTCAGAGCTGCTGTCATTTTAGACCCGTGATGCTGAAGG 53
1204 yLeuPheGlyLeuHisLysLysValIleAsnMetValHisAsnLeuLeuS 1221
|||||
54 ACTGTTTGAGATTGCATTAAGAAAGTCATCAATATGTGACAAATTTACTCT 103
1221 erSerHisAspSerAspProArgTyrSerAspProGlnIleLysAlaArg 1237
|||||
104 CCAGTCACGACTCAGACCCGCGTACTCTGACCCCTCAGATAAAGGCTCGA 153
1238 ValAlaMetLeuTyrLeuProIleuIleGlyIleIleMetGluThrValPr 1254
|||||
154 GTGGCCATGTGTATCTACCTCTGATGTGATATCATGGAAGTGTACC 203
1254 oGlnLeuTyrAspPheThrGluThrHisAsnGlnArgGlyArgProIleC 1271
|||||
204 TCAGCTGTATGATTTTACAGAAACTCACAATCAACGAGAGAACCAATTT 253
1271 ysiLeaIaThrAspAspTyrGluSerGluSerGlySerMetIleSerGln 1287
|||||
254 GTATAGCCACTGATGATTATGAAGTGAGAGCGGAAGTATGATAAGCCAG 303
1288 ThrValAlaMetAlaIleAlaGlyThrSerValProGlnLeuThrArgPr 1304
|||||

304 ACCGTTGCCATGGCAATCGCAGGACATCGGTCCCTCAACTAACAGGCC 353
1304 oGlySerPheLeuLeuThrSerThrSerGlyArgGlnHisThrThrPheS 1321
|||||
354 TGGCAGTTTCCTCCTCACGTCACAGAGTGGCAGGACACACACTACCTTT 403
1321 erAlaGluSerSerArgSerLeuLeuIleCysLeuLeuTrpValLeuLys 1337
|||||
404 CAGCAGAAATCAAGTCGAAGCCTTTGATCTGTCTACTTGGGTTCTCAAA 453
1338 AsnAlaAspGluThrValLeuGlnLysTrpPheThrAspLeuSerValle 1354
|||||
454 AATGCAGATGAACAGTCTCTACAGAAAGTGTTCACAGATCTCTCAGTCTT 503
1354 uGlnLeuAsnArgLeuLeuAspLeuLeuTyrLeuCysValSerCysPheG 1371
|||||
504 GCAGCTAAACCGGCTATTTAGATCTGCTTTATCTCTGTGTGCTTTGG 553
1371 lUTyrLysGlyLysLysValPheGluArgMetAsnSerLeuThrPheLys 1387
|||||
554 AGTATAAAGGGAAGAAAGTGTGTAACGAATGAATAGCTTGACCTTTAAG 603
1388 LysSerLysAspMetArgAlaLysLeuGluAlaIleLeuGlySerI 1404
|||||
604 AAATCAAAAGACATGAGAGCAAGCTTGAAGAAAGCTATCTTGGAGCAT 653
1404 eGlyAlaArgGlnGluMetValArgArgSerArgGlyGlnLeuGluArgS 1421
|||||
654 AGTGCCAGGCAAGAAATGTGTACGGCGAAGCCGAGAGACAGCTCGAGAGA 703
1421 erProSerGlySerAlaPheGlySerGlnGluAsnLeuArgTrpArgLys 1437
|||||
704 GCCCATCTCGAAGAGTGCCCTTTGAGAGTCAAGAAATTTGAGGTGAGGAAA 753
1438 AspMetThrHisTrpArgGlnAsnThrGluLysLeuAspLysSerArgAl 1454
|||||
754 GATATGACTCTACTGGCGTCAAAACACACAGAGAAGCTTGACAAATCAAGAGC 803
1454 aGluIleGlnHisGluAlaLeuIleAspGlyAsnLeuAlaThrGluAla 1471
|||||
804 AGAGATTGAACACGAAGCAGCTGATGTGGAACCTGCGCTACAGAACCAA 853
1471 snLeuIleIleLeuAspThrLeuGluIleValValGlnThrValSerVal 1487
|||||
854 ACCTAATCATTTAGATACATTAAGATTGTTGTTTCAGACCGTTCTGTGA 903
1488 ThrGluSerLysGluSerIleLeuGlyGlyValLeuLysValLeuLeuHi 1504
|||||
904 ACGGAATCCAAAGAGAGACATTTCTGTGGAGTGCTAAAGAGTCTACTACA 953
1504 sSerMetAlaCysAsnGlnSerAlaValTyrLeuGlnHisCysPheAlaT 1521
|||||
954 CAGCATGGCCCTGTAACCAAGTGCAGTTTATCTACACACACTGTTTGCTA 1003
1521 hrGlnArgAlaLeuValSerLysPheProGluLeuLeuPheGluGlu 1537
|||||
1004 CACAGAGAGCCTTGTTTCAAAAGTTTCTGAACTCTTATTTGAAGAAGAG 1053
1538 ThrGluGlnCysAlaAspLeuCysLeuArgLeuLeuArgHisCysSerSe 1554
|||||
1054 ACAGAGCAGTGTCTGATTTATGCTCAGGCTTCTCCGACACACTGTAGCAG 1103
1554 rSerIleGlyThrIleArgSerHisProSerAlaSerLeuTyrLeuLeuM 1571
|||||
1104 TAGCATCGGTACATAACGGTCACACCCAGTGCCCTTACCTACTAA 1153
1571 eTArgGlnAsnPheGluIleGlyAsnAsnPheAlaArgValLysMetGln 1587
|||||
1154 TGAGGCAAAACTTTGAGATTGGGAATACCTTGCCAGGGTTAAATGCAG 1203
1588 ValPrometSerLeuSerSerLeuValGlyThrSerGlnAsnPheAsnGl 1604
|||||
1204 GTACCAATGTCTACTATCCTCCTTGGTGGCACAATCTCAGAAATTTTAATGA 1253

1604 uGluPheLeuArgArgSerLeuLysThrIleLeuThrTyrAlaGluGluA 1621
|||||
1254 AGAATCTTAGAGCGTTCTCTAAAGACTATATTGACATATGCTGAGAAG 1303
1621 spleuGluLeuArgGluThrThrPheProAspGlnValGlnAspLeuVal 1637
|||||
1304 ATCTGGAATGAGGGAACAACACATTTCCCTGATCAGGTCAGGATCTGGTT 1353
1638 PheAsnLeuHisMetIleLeuSerAspThrValLysMetLysGluHisG1 1654
|||||
1354 TTCAATCTCCATATGATTTCTTCTGATACTGTGAAAATGAAGGAACACCA 1403
1654 nGluAspProGluMetLeuIleAspLeuMetTyrArgIleAlaLysGlyT 1671
|||||
1404 GGAGGATCCTGAATGTTGATTGATCTAATGTACAGAATTCGCAAGGGTT 1453
1671 yrgLInThrSerProGlu..ArgLeuThrTrpLeuGlnAsnMetAlaGly 1686
|||||
1454 ACCGAGACCTCTCCAGATCTGCGATTCGACTGCTGTCAGAACATGCGACGC 1503
1687 LysHisSerGluArgSerAsnHisAlaGluAlaAlaGlnCysLeuValHi 1703
|||||
1504 AAGCACTCAGAACGAAACAATCATGCTGAAGCTGCACAGTGTCTAGTCCA 1553
1703 sSerAlaIaLeuValAlaGluTyrLeuSerMetLeuGluAspArgLysT 1720
|||||
1554 CTCAGCAGCACTTGTGCTGTAATATTGAGCATGTCTGGAGACCGGAAT 1603
1720 yrrLeuProValGlyCysValThrPheGlnAsnIleSerSerAsnValLeu 1736
|||||
1604 ATCTTCCTGCTGGGATGTGTAACATTTTCAGAAATATTTCAATCTAATGTTTTA 1653
1737 GluGluSerAlaValSerAspAspValValSerProAspGluGlyI1 1753
|||||
1654 GAAGAATCTGCGGCTCTCAGATGATGTGTATCTCCAGATGAAGAAAGGTAT 1703
1753 eCysSerGlyLysTyrPheThrGluSerGlyLeuValGlyLeuGluG 1770
|||||
1704 CTGCTCTGGAATAACTTTACTGAGTCAAGACTGTGGGATTACTGGAAC 1753
1770 InAlaIaIaIaSerPheSerMetAlaGlyMetTyrGluAlaValAsnGlu 1786
|||||
1754 AAGCAGCTGCTTCTCTCTATGCGTGCATGTATGAGCAGCTTAATGAA 1803
1787 ValTyrLysValLeuIleProIleHisGluAlaAsnArgAspAlaLysLy 1803
|||||
1804 GTTTACAAAGTACTTATTTCCATTTCAAGAAGCTAATCGGATGCCAAAGAA 1853
1803 sLeuSerThrIleHisGlyLysLeuGlnGluAlaPheSerLysIleValH 1820
|||||
1854 ACTATCCACAATTCATGTTAACTTCAAGAAAGCATTCAGCAAAATTGTTTC 1903
1820 isGlnSerThrGlyTrrpGluArgMetPheGlyThrTyrPheArgValGly 1836
|||||
1904 ATCAGAGTACTGGCTGGAGCGGATGTTTGGCACCTAATTTTCGTGTGGT 1953
1837 PheTyrGlyThrLysPheGlyAspLeuAspGluGlnGluPheValTyrLy 1853
|||||
1954 TTTTATGGAACCAAGTTCGGGATTTGGATGAACAAGAAATTGTTTACAA 2003
1853 sGluProAlaIleThrLysLeuAlaGluIleSerHisArgLeuGluGlyP 1870
|||||
2004 GGAGCCTGCATAATACCAAACTTGACAGATATCTCAGAGATTGAGGAGAT 2053
1870 heTyrGlyGluArgPheGlyGluAspValValGluValIleLysAspSer 1886
|||||
2054 TTTACGAGAGAAAGATTGGAGAGGATGTGTTGAAGTAATCAAGACTCT 2103
1887 AsnProValAspLysCysLysLeuAspProAsnLysAlaTyrIleGlnI1 1903
|||||
2104 AATCCTGTAGACAAGTGTAAATTAGATCCTTAACAAGGCATATATTTCAGAT 2153

1903 ehrrTyrValGluProTyrPheAspThrTyrGluMetLysAspArgIleT 1920
|||||
2154 TACCTATGTGAGCCCACTACTTTGACACATATGAGATGAAGACAGATCA 2203
1920 hrTyrPheAspLysAsnTyrAsnLeuArgArgPheMetTyrCysThrPro 1936
|||||
2204 CCTATTTGCAGAAAATTTACAATCTTCGTCGATTCATGTACTGTACACCC 2253
1937 PheThrLeuAspGlyArgAlaHisGlyGluLeuHisGluGlnPheLysAr 1953
|||||
2254 TTTACTTTAGATGGCCGTGCCCATGGGGAAGCTTCATGAACAATTCAAAAG 2303
1953 gLysThrIleLeuThrThrSerHisAlaPheProTyrIleLysThrArgV 1970
|||||
2304 GAAGACCATTTCTGACTACGTCTCATGCTTTCCTTATATTAATAACAGGG 2353
1970 aIAsnValThrHisLysGluGluIleIleLeuThrProIleGluValAla 1986
|||||
2354 TCAATGTCACTCATTAAGAAGAGATCATCTTAACACCAATTGAAGTTGCT 2403
1987 ileGluAspMetGlnLysLysThrGlnGluLeuAlaPheAlaThrHisG1 2003
|||||
2404 ATTGAGGACATGCAGAAAAGACACAGAGTTGGCATTTGCAACACATCA 2453
2003 nAspProAlaAspProLysMetLeuGlnMetValLeuGlnGlySerValG 2020
|||||
2454 GGATCCCGCAGACCCCAAAATGCTTCAGATGTACTCCAGGGATCTGTAG 2503
2020 lyThrThrValAsnGlnGlyProLeuGluValAlaGlnValPheLeuSer 2036
|||||
2504 GCACACAGAGTGAATCAGGGGCCCTTTGGAAGTTGCCACAGGTTTTCTGTCT 2553
2037 GluIleProSerAspProLysLeuPheArgHisHisAsnLysLeuArgle 2053
|||||
2554 GAAATACCTAGTGAACCCAAAGCTCTTCAGACATCATATAAATGCGACT 2603
2053 uCysPheLysAspPheThrLysArgCysGluAspAlaLeuArgLysAsnL 2070
|||||
2604 CTGCTTTAAAGATTTTACTAAAGAGGTGAAGATGCCTTAAGAAAATA 2653
2070 ysSerLeuIleGlyProValGlnLysGluTyrGlnArgGluLeuGlyLys 2086
|||||
2654 AGAGCTTAATTCGCGCGGTTCAAAAAGAGATATCAAAAGGAATTCGGGAAA 2703
2087 leuSerSerPro 2090
|||||
2704 CTATCTTCGCT 2715

seq_name: gb_pr:AK024436

seq_documentation_block: 4577 bp mRNA linear PRI 29-SEP-2000
LOCUS AK024436
DEFINITION Homo sapiens mRNA for FLJ00026 protein, partial cds.
ACCESSION AK024436
VERSION AK024436.1 GI:10440380
KEYWORDS fis (full insert sequence).
SOURCE Homo sapiens adult spleen cDNA to mRNA, clone:as00026.
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 4577)
AUTHORS Ohara,O., Nagase,T., Kikuno,R. and Okumura,K.
TITLE The nucleotide sequence of a long cDNA clone isolated from human spleen

JOURNAL Published Only in DataBase (2000) In press
REFERENCE 2 (bases 1 to 4577)
AUTHORS Ohara,O., Nagase,T., Kikuno,R. and Okumura,K.
TITLE Direct Submission

Submitted (24-AUG-2000) Osamu Ohara, Kazusa DNA Research Institute,
Department of Human Gene Research; 1532-3, Yana, Kisarazu, Chiba
292-0812, Japan (E-mail:cdnainfo@kazusa.or.jp,
URL:htp://www.kazusa.or.jp/NEDO, Tel:81-438-52-3913,
Fax:81-438-52-3914)

1081 CTTGCCCTATAAGCAGTACAAACATGCTGAACGGGACACTACTCGCAACC 1130
1328 euleuIleCysleuLeuTrpValleuLysasnAlaAspGluThrValleu 1344
1131 TCATGATCTGCTCTCTCGATCATGAAAAATGCTGATCAGAGCCCTCATT 1180
1345 GlnLysTrpPheThrAspLeuSerValleuGlnLeuAsnArgLeuLeuAs 1361
1181 AGGAAGTGGAATGCTGACCTGCCATCAACGCAGCTCAACAGATTTTGA 1230
1361 pLeuLeuTyrlLeuCysValSerCysPheGluTyrlLysGlyLysValP 1378
1231 TCTACTTTTCATCTGCTGTGTTATGTTTGAGTATTAAGGAAAAACAGAGT 1280
1378 heGluArgMetaSerLeuThrPheLysLysSerLysAspMetaArgAla 1394
1281 CTGACAAAGTCAGTACCCAAAGTCCCTGCAGAACTCAAGGATGTCAAGCC 1330
1395 LysLeuGluGluAlaIleLeuGlySerIleGlyAlaArgGlnGluMetVa 1411
1331 CGGCTGGAAGAGGCTTTGCTCCGTGGGGAAGGGGCCAGAGGGGAGATGAT 1380
1411 LArgArgSerArgGlyGlnLeuGluArgSerProSerGlySerAlaPhe. 1427
1381 GCGCCGC.....CGGGCTCCAGGGGAACGACCGATTTTC 1412
1428 ..GlySerGlnGluAsnLeuArgTrpArgLysAspMetThrHisTrpArg 1443
1413 CAGGCCCTAAATGAAAAATTGAGATGGAAGAAAGAGACACACATTTGGCGG 1462
1444 GlnAsnThrGluLysLeuAspLysSerArgAlaGluIleGlnHisGluAl 1460
1463 CAAGCTAATGAGAAAGCTAGATAAAACAAGCCGAGTTAGATCAAGAAAGC 1512
1460 aLeuIleAspGlyAsnLeuAlaThrGluAlaAsnLeuIleLeuAspT 1477
1513 CTTGATCAGTGGCAATCTGGCTACAGAACACATTTAATCATCTCGATA 1562
1477 hrLeuGluIleValIleGlnThrValSerValThrGluSerLysGluSer 1493
1563 TGCAGGAAGAAACATTTATCCAGCGAGCTCGGCTCTGACTGTAAAGACAGC 1612
1494 IleLeuGlyGlyValleuLysValleuLeuHisSerMetaIacysAsnG 1510
1613 CTGCTGGAGGTGTCTGAGGGTGTGTTGAATTTCTCTGACTGTGATCA 1662
1510 nSerAlaValTyrlLeuGlnHisCysPheAlaThrGlnArgAlaLeuValS 1527
1663 GAGTACCACCTACTGACTCACTGTTTGCAACACTCCGTCTCTCATCG 1712
1527 erLysPheProGluLeuLeuPheGluGluGluThrGluGlnCysAlaAsp 1543
1713 CCAAGTTTGAGACTTACTCTTCGAAGAGGAGGTGAACAGTGTTCGAC 1762
1544 LeuCysLeuArgLeuLeuArgHisCysSerSerSerIleGlyThrIleAr 1560
1763 CTATGTACCAAGTCTCTGCACCACTGCAGCAGCAGCATGATGTACCCG 1812
1560 gSerHisProSerAlaSerLeuTyrlLeuLeuMetaArgGlnAsnPheGlu 1577
1813 GAGCCAAAGCCTGTGCCACCTTTAACCCTCATGAGGTTTCAGTTTGGAG 1862
1577 leGlyAsnAsnPheAlaArgValLysMetGlnValPrometSerLeuSer 1593
1863 CCACCAGTAATTTTGCAAGAGTAAGATGCAAGTAACCATGTCCCTGGCA 1912
1594 SerLeuValGlyThrSerGlnAsnPheAsnGluLupheLeuArgArgSe 1610
1913 TCTTTGGTGGGAAGAGCACCAGACTTTAATGAAGAGCACCTGAGAAAGATC 1962
1610 rLeuLysThrIleLeuThrTyrlAlaGluGluAspLeuGluLeuArgGluT 1627

1963 CTTGAGGACAATTTTGCCCTATTTCAGAAAGAGACACAGCCATGCAGATGA 2012
1627 hrThrPheProAspGlnValGlnAspLeuValPheAsnLeuHisMetIle 1643
2013 CTCCTTTTCCACCAGAGGTGAGGAAGTCTCTGTATCTGATAGCATC 2062
1644 LeuSerAspThrValLysMetLysGluHisGlnGluAspProGluMetLe 1660
2063 TTATATGACACAGTGAATAAGGGAATTTTCAGGAAGATCCTGAGATGCT 2112
1660 uIleAspLeuMetTyrlArgIleAlaLysGlyTyrlGlnThrSerProGlu. 1676
2113 TATGATCTCATGTACAGAAATTGCCAAGAGTTACCAAGGCATCTCTGATC 2162
1677 ..ArgLeuThrTrpLeuGlnAsnMetaIaGlyLysHisSerGluArgSer 1692
2163 TCGGCTGACCTGGCTCCAGAAACATGGCAGAGAAACACACCAAGAAAG 2212
1693 AsnHisAlaGluAlaIleGlnCysLeuValHisSerAlaAlaLeuValAl 1709
2213 TGCTTACACGAGGCTGCCATGTGCCGTGTCAGCCCGCTGCGTTAGTGGC 2262
1709 aGluTyrlLeuSerMetLeuGluAspArgLysTyrlLeuProValGlyCysV 1726
2263 TGAGTATCTGAGCATGCTGGAGGACACACAGCTACTGCCCGTGGGCAGTG 2312
1726 alThrPheGlnAsnIleSerSerAsnValleuGluGluSerAlaValSer 1742
2313 TCAGCTTCCAGAAATATTCTTCCAAATGCTCGGAGGAGTCTGTGTCTCT 2362
1743 AspAspValValSerProAspGluGluGlyIleCysSerGlyLysTyrlP 1759
2363 GAGGACACCCCTGTCACTTGACGAGAGATGGGTGTGGCAGGCCAGTACTT 2412
1759 eThrGluSerGlyLeuValGlyLeuLeuGluGlnAlaAlaIleSerPheS 1776
2413 CACCGAGAGTGGCTGTGTAGGCCCTCTGAGCAGGCCGAGAGCTGTCA 2462
1776 erMetaIaGluMetTyrlGluAlaValaIleAsnGluValTyrlLysValleuIle 1792
2463 GCACGGGAGGCTGTATGAGACAGATTAAATGAGGTCTACAAGCTGTGATC 2512
1793 ProIleHisGluAlaAsnArgAspAlaLysLysLeuSerThrIleHisG 1809
2513 CCCATCTTAGAAGCGCATCGAAGATTTCCGGAAGCTGACACTCACTCACAG 2562
1809 yLysLeuGlnGluAlaPheSerLysIleValHisGlnSerThrGlyTrpG 1826
2563 CAAGCTGCAGAGAGCCCTTCGACAGCATCGTTTAACAAGAT.....CATA 2606
1826 LuArgMetPheGlyThrTyrlPheArgValGlyPheTyrlGlyThrLysPhe 1842
2607 AGAAGATGTTTGAACCTACTCTCCGAGTTGGTTCTTTGGATCCAAATTT 2656
1843 GlyAspLeuAspGluGlnGluPheValTyrlLysGluProAlaIlePheTr 1859
2657 GGGGATTTGGATGAACAGGAATTTGTACAAAGAGCCCTGCAATTACCAA 2706
1859 sLeuAlaGluIleSerHisArgLeuGluGlyPheTyrlGlyGluArgPheG 1876
2707 GCTTCCCTGAGATCTCACATAGACTAGAGCAATTTATGTGTCATATTTTG 2756
1876 lyGluAspValValGluValIleLysAspSerAsnProValAspLysCys 1892
2757 GTGCAGAAATTTGTGAAGTGAATTAAGACTCCACTCCTGTGCAAAAACC 2806
1893 LysLeuAspProAsnLysAlaTyrlIleGlnIleThrTyrlValGluProTy 1909
2807 AAGTTGATCTCTAACAAAGCCCTACATACAGATCACTTTTGTGAGAGCCCTA 2856
1909 rPheAspThrTyrlGluMetLysAspArgIleThrTyrlPheAspLysAsnT 1926
2857 CTTTGATGAGTATGAGATGAAGAGACAGGGTCACATACTTTTGAGAAAGATT 2906

1926 yfAsnLeuArGArpHeMeTYrCysThrProPheThrLeuAspGLyArg 1942
:::|||||
2907 TCAACCTCCGGAGGTTCATGTACACCACCCGTCACCCGTGAGGGCGCG 2956
1943 AlAHISgIyGluLeuHISgIuGlnPheLYsArgLYsThrIleLeuThrTh 1959
:::|||||
2957 CCTCGGGAGAGCTGCATGAGCAGTACAGAAAGAACACAGTCTGACCAC 3006
1959 rSerHISAlaPheProTYrIleLYsThrArgValAsnValThrHISLYSg 1976
| |||||
3007 TATGCACGCCCTTCCCTACATCAAGACCAAGATCAGCGTCATCCAGAAG 3056
1976 lUGluIleIleLeuThrProIleGluValAlaIleGluAspMetGlnLYs 1992
|||||
3057 AGGAGTTGTTTGTGACACCGATGAAGTTGCCATGAGACATGAAGAAG 3106
1993 LYsThrgInGluLeuAlaPheAlaThrHISgInAspProAlaAspProLY 2009
||||| ::||| |
3107 AAGACCTGCAGTTAGCAGTTGCCATTAACCAAGAGCGCCTGATGCAAA 3156
2009 sMetLeuGlnMetValLeuGlnLYsSerValGLyThrThrValAsnGlnG 2026
|||||
3157 GATGCTTCAGATGGTGTGCAAGGCTGTGGGAGCTACTGTAATCAGG 3206
2026 lYProLeuGluValAlaGlnValPheLeuSerGluIleProSerAspPro 2042
|||||
3207 GACCACGTGAAGTAGCCCAAGTTTGGCTGAATTCCTGCTGATCCA 3256
2043 LYsLeuPheArGHisHISAsnLYsLeuArGLyLeuCysPheLYsAspPheTh 2059
|||||
3257 AAACCTATCGACATCACACAAGTTGAGGTTATGCTTTAAGGAATTTCAT 3306
2059 rLYsArgCysGluAspAlaLeuArGLysAsnLYsSerLeuIleGLyProV 2076
||||| ::||| |
3307 CATGAGATGTGTGAAGCTGTAGAGAAAAACAAGCGTCTCATCACGGCAG 3356
2076 aGlnLYsGluTYrGlnArGLyLeuGLyLYs 2086
|||||
3357 ACCAGAGGGAATATCAGCAGGAACTCAAAAAAG 3388

SQ Sequence 6828 BP; 2082 A; 1384 C; 1432 G; 1930 T; 0 other;

alignment_scores:

Quality: 10815.00 Length: 2090
Ratio: 5.175 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-737-246-2 x AAH41934 ..

Align seg 1/1 to: AAH41934 from: 1 to: 6828

1 MetAlaGluArgArgAlaPheAlaGlnLysIleSerArgThrValAlaAl 17
|||||
22 ATGGCCGAGCGCCGCCCTTCGCCAGAGATCAGCAGAACGGTGCAGC 71
17 agluValArgLysGlnIleSerGlyIntyrSerGlySerProGlnLeuL 34
|||||
72 CGAAGTTAGAGACAGATCTCCGACACATATAGTGGTCTCCCAACTGC 121
34 eulysAsnLeuAsnIleValGlyAsnIleSerHisIleThrThrValPro 50
|||||
122 TCAAAAACCTTAATATTGTTGGCAATATATCCCATCACACCACAGTGCCC 171
51 LeuThrGluAlaValAlaSPProValAspLeuGluAspTyrLeuIleThrHi 67
|||||
172 CTACCGAAGCAGATGATCCAGTGGATTGGAAGATTACCTCATTTACTCA 221
67 sProLeuAlaValAspSerGlyProLeuArgAspLeuIleGluPheProp 84
|||||
222 TCCTTTGGCTGTGCAATTCTGGGCCCTTACGGGATTTGATTGAATTTCCCTC 271
84 roAspAspIleGluValValIleTyrSerProArgAspCysArgThrLeuVal 100
|||||
272 CAGATGATATGAAAGTTGTTATAGTCCCTCGGAGACTGCAGACACTCTGTT 321
101 SerAlaValProGluGluSerGluMetAspProHisValArgAspCysIle 117
|||||
322 TCAGCTGTACCTGAAGAAGTGAATGATGCCACATGTTAGAGACTGTAT 371
117 eArgSerTyrThrGluAspTPRAlaIleValIleArgLysTyrHisLysL 134
|||||
372 AAGAAGTTATACAGAAGACTGGGCAATTGTCAATCAGAAAATATCATTAAT 421
134 euGlyThrGlyPheAsnProAsnThrLeuAspLysGlnLysGluArgGln 150
|||||
422 TGGGAACAGGATTTAATCCCAATACATTAGATTAACAGAAAGCAAA 471
151 LysGlyLeuProLysGlnValPheGluSerAspGluAlaProAspGlyAs 167
|||||
472 AAAGTTTGCCAAACAAGTTTGTGAATCTGATGAAGCTCCAGATGGCAA 521
167 nSerTyrGlnAspAspGlnAspAspLeuLysArgArgSerMetSerIleA 184
|||||
522 CAGCTACCAGAGATGATCAAGATGACCTTAAGACGTTCATGTCATAG 571
184 sPAspThrProArgGlySerTyrPalacysserIlePheAspLeuLysAsn 200
|||||
572 ATGATACCCCAAGGGGTAGCTGGGCTGTAGTATCTTTGACTTTGAAAAAT 621
201 SerLeuProAspAlaLeuLeuProAsnLeuLeuAspArgThrProAsnGln 217
|||||
622 TCACCTTCCTGATGCTTTGCTTCCCAATTTACTTGATCGAACTCCAAATGA 671
217 uGluIleAspArgGlnAsnAspAspGlnArgLysSerAsnArgHisLysG 234
|||||
672 AGAAATAGACCGTCAGAATGATGACCAAGGAATCAAAACCGTCACAAG 721
234 IuLeuPheAlaLeuHisProSerProAspGluGluGluProIleGluArg 250
|||||
722 AACTTTTGTGCTTGCATCCATCACCAAGATGAGGAAGAACCAATAGAACGG 771

251 LeuSerValProAspIleProLysGlnHisPheGlyGlnArgLeuLeuVa 267
|||||
772 CTTAGTGTCTCCTGATATACCCAAAGACATTTGGTCAAGACTTCTGT 821
267 lLysCysLeuSerLeuLysPheGluIleGluIleGluProIlePheAlas 284
|||||
822 AAAATGCTTATCACCTCAGATTGAAATTGAAATGAAACCATTTTTCGA 871
284 erLeuAlaLeuTyrAspValLysGluLysLysLysIleSerGluAsnPhe 300
|||||
872 GTTTGGCTTTATATGATGTCAAGGAAAAAGAAAGATTTCAGAAAACTTT 921
301 TyrPheAspLeuAsnSerGluGlnMetLysGlyLeuLeuArgProHisVa 317
|||||
922 TATTTTGACCTTAATCTGAGCAGATGAAAGGTTGTTACGTCACATGT 971
317 lProProAlaAlaIleThrThrLeuAlaArgSerAlaIlePheSerIleT 334
|||||
972 ACCACCTGCTGCCATTTACTACCCCTGGCAAGATCAGCAATTTTTCATCA 1021
334 hrTyrProSerGlnAspValPheLeuValIleLysLeuGluLysValLeu 350
|||||
1022 CTTATCCTTCCCAAGATGTTTTCTTGTAATTAAGCTAGAAAAAGTCTTA 1071
351 GlnGlnGlyAspIleGlyGluCysAlaGluProTyrMetIlePheLysGln 367
|||||
1072 CAGCAAGGAGACATTTGAGAGTGTGCAGAACCATATATGATTTTCAACA 1121
367 uAlaAspAlaThrLysAsnLysGluLysLeuGluLysLeuLysSerGlnA 384
|||||
1122 AGCAGATGCCACCAAGATTAAGAAAACTGGAGAAACTGAAGAGTCAAG 1171
384 laAspGlnPheCysGlnArgLeuGlyLysTyrArgMetProPheAlaTyr 400
|||||
1172 CAGATCAGTTTTGCCAAGACTTGGGAATATCGCATGCCCTTTGCTGG 1221
401 ThrAlaIleHisLeuMetAsnIleValSerSerAlaGlySerLeuGluAr 417
|||||
1222 ACTGCAATCCATTTAATGAATATTGTTAGACAGTGCCTGGAGTTTGGAAAG 1271
417 gasPserThrGluValGluIleSerThrGlyGluArgLysGlySerTrps 434
|||||
1272 AGATTCTACAGAGATGAATAATCAGTACTGAGAACGAAAGGCTGTGCT 1321
434 erGluArgArgAsnSerSerIleValGlyArgArgSerLeuGluArgThr 450
|||||
1322 CAGAGAGGAGGAATTTAGTATTGTTGGCAGACGATCACTTGAAGAGACA 1371
451 ThrSerGlyAspAspAlaCysAsnLeuThrSerPheArgProAlaThrLe 467
|||||
1372 ACAAGTGAGATGATGCTTGTAACTTGACGAGCTTTCAGACCAGCTACTCT 1421
467 uThrValThrAsnPhePheLysGlnGluGlyAspArgLeuSerAspGluA 484
|||||
1422 CACAGTGACAAATTTTAAAGCAGGAAGGAGACCGCTTAAGTGATGAG 1471
484 sPLeuTyrLysPheLeuAlaAspMetArgArgProSerSerValLeuArg 500
|||||
1472 ATCTCTACAATAATTCCTGCTGATATGAGAGGCCATCTTCTGTCTACGG 1521
1522 ArgLeuArgProIleThrAlaGlnLeuLysIleAspIleSerProAlaPr 517
|||||
1522 CGACTAAGACCTATTTACAGCTCAGCTCAAGATAGACATTTCTCCCGCACC 1571
517 oGluAsnProHisTyrCysLeuThrProGluLeuLeuGlnValLysLeuT 534
|||||
1572 TGAATAATCCCATTTATTTGCCCTAACTCCGAGCTGCTTCAAGTGAAGCTTT 1621
534 yrProAspSerArgValArgProThrArgGluIleLeuGluPheProAla 550
|||||
1622 ACCCTGACAGTAGAGTTAGACCTACAGAGAAATCTTAGAGTTTCCGCA 1671
551 ArgAspValTyrValProAsnThrThrTyrArgAsnLeuLeuTyrIleTy 567

|||||
1672 AGGAGTGTATGTTCCAAACACTACTTACAGAAATCTCTCATATATA 1721
567 rProGlnSerLeuAsnPhelaAsnArgGlnGlySerAlaArgAsnIleT 584
|||||
1722 CCGTCAGAGTCTTAATTTGCCAATCGTCAGAGTCTCGTAGAATATATA 1771
584 hrValLysValGlnPheMetTyrGlyGluAspProSerAsnAlaMetPro 600
|||||
1772 CAGTGAAGTCCAGTATTATGATGAGAGGATCCAGCAATGCCATGGCG 1821
601 ValIlePheGlyLysSerSerGluPheSerLysGluAlaTyrTh 617
|||||
1822 GTAATCTTTGGTAAATCTAGCTGTCAGAAATTTCAAGGAAGCCTATAC 1871
617 rAlaValAlaTyrHisAsnArgSerProAspPheHisGluIleLysV 634
|||||
1872 AGCCGTAGTATATCATAAACAGGTCCTGATTTTCATGAAGAAATCAAG 1921
634 alLysLeuProAlaThrLeuThrAspHisHisIleLeuPheThrPhe 650
|||||
1922 TTAAGCTTCCTGCTACTTTAAGTCAGCATCATCTGCTTTTACTTT 1971
651 TyrHisValSerCysGlnGlnLysGlnAsnThrProLeuGluThrProVa 667
|||||
1972 TATCATGTTAGTGTGTCACAAAAACAATACTCTCTTGAACACACAGT 2021
667 IGLYTYRTHRTTRIPLEPROMETLEUGLNASNGLYARGLEULYSTHRLYG 684
|||||
2022 TGGATATACATGGATACCAATGCTTCAGAATGGACGGTTGAAGACTGGCC 2071
684 lnpheCysLeuProValSerLeuGluLysProProGlnAlaTyrSerVal 700
|||||
2072 AGTTTGGCTTGGCCAGTCTCATTTGAAAAACCACACAGGCTTATTTCTGTA 2121
701 LeuSerProGluValProLeuProGlyMetLysTrpValAspAsnHisLy 717
|||||
2122 CTGTCTCCAGAGGTTCCTCTACCTGGCATGAATGGTAGATAATCACAA 2171
717 sGlyValPheAsnValGluValAlaValSerSerIleHisThrGlnA 734
|||||
2172 AGGTGTTTAAATGTTGAAGTGTGCTGTTTCGTATCCATACACAAG 2221
734 spProTyrLeuAspLysPhePheAlaLeuValAsnAlaLeuAspGluHis 750
|||||
2222 ATCCTTATCTTGACAAATTTTGTCTCTGTCATGCTCTGGATGAACAC 2271
751 LeuPheProValArgIleGlyAspMetArgIleMetGluAsnAsnLeuGl 767
|||||
2272 CTGTCCAGTCCGAATTTGGGACATGCCGAATCATGGAATAACTTAGA 2321
767 uAsnGluLeuLysSerSerIleSerAlaLeuAsnSerSerGlnLeuGluP 784
|||||
2322 AAATGAATGGAAGCAGATATTTCAGCACTGAATTCATCCCACTGGAAC 2371
784 roValValArgPheLeuHisIleLeuLeuAspLysLeuIleLeuLeuVal 800
|||||
2372 CAGTGGTCGATTTCTTCATCTTCGCTAGATAAACTGATACTTTAGTT 2421
801 IleArgProProValIleAlaGlyGlnIleValAsnLeuGlyGlnAlaSe 817
|||||
2422 ATTAGACCTCCTGTCATTTGCTGGCCAAATAGTTAACCTAGGTCAAGCATC 2471
817 rpheGluAlaMetAlaSerIleIleAsnArgLeuHisIleLysAsnLeuGluG 834
|||||
2472 TTTGAAGCCATGGCATCAATATATAATCGACTTCACAAAAAACTTGAAG 2521
834 lYAsnHisAspGlnHisGlyArgAsnSerLeuLeuAlaSerTyrIleHis 850
|||||
2522 GAAATCATGACCAAGCATGGCAGAAACAGCCTTCTGCATCATATATTCAT 2571
851 TyrValPheArgLeuProAsnThrTyrProAsnSerSerSerProGlyPr 867
|||||

2572 TATGTTTCCGCCCTACCAATACTTACCCTAATTTCATCATCACAGGTCC 2621
867 oGlyGlyLeuGlyGlySerValHisTyrAlaThrMetAlaArgSerAlaV 884
|||||
2622 TGGGGTTTGGGAGGATCAGTGCATTTATGCCACAATGGCTAGATCTGCCG 2671
884 aLArgProAlaSerLeuAsnLeuAsnArgSerArgSerLeuSerAsnSer 900
|||||
2672 TGAGACCTGCAAGCCTTAATTTAAATCGTTCTCGAAGCCTTAGTAATAGC 2721
901 AsnProAspIleSerGlyThrProThrSerProAspAspGluValArgSe 917
|||||
2722 AATCCAGATATATCTGGGACTCCACAGTCACCAAGATGAAGTTCCGATC 2771
917 rIleIleGlySerLysGlyLeuAspArgSerAsnSerTrpValAsnThrg 934
|||||
2772 AATCATCGGAGTAAGGTTTAGATCGCTCCAATTCCGCGTTAACACTG 2821
934 lYGLYPROLYSAlaAlaProTrpGlySerAsnProSerProSerAlaGlu 950
|||||
2822 GTGCTCCAAAGCTGCCCCATGGGATGCCAACCCAGTCCAAAGTGCAGAA 2871
951 SerThrGlnAlaMetAspArgSerCysAsnArgMetSerSerHisThrGl 967
|||||
2872 TCAACACAGGCTATGATCGAAGTTGATATCGTATGTTCTTCGCACAGAGA 2921
967 uThrSerSerPheLeuGlnThrLeuThrGlyArgLeuProThrLysLysL 984
|||||
2922 GACGTCAAGTTTCTTACAAACATTAAACGGGACGCTTACCACAACTAAAGC 2971
984 eupheHisGluGluLeuAlaLeuGlnTrpValValCysSerGlySerVal 1000
|||||
2972 TTTTTCACGAGAGAGTGGCTTTCAGATGGGTTGTTGCAGTGGCAGCGTT 3021
1001 ArgGluSerAlaLeuGlnGlnAlaTrpPhePheGluLeuMetValLy 1017
|||||
3022 CGGGAATCAGCTTTGCAACAAGCCGTGTTCTTTTGAATTAATGGTAA 3071
1017 sSerMetValHisHisLeuTyrPheAsnAspLysLeuGluAlaProArgL 1034
|||||
3072 GAGCATGGTGACCAATTATACTTTAATGATAACTTGAGGCTCCAGGA 3121
1034 ysSerArgPheProGluArgPheMetAspIleAlaAlaLeuValSer 1050
|||||
3122 AAAGTCGTTTCCAGAACGTTTCATGATGACATTTGCAGCTCTGTGCAGC 3171
1051 ThrIleAlaSerAspIleValSerArgPheGlnLysAspThrGluMetVa 1067
|||||
3172 ACGATTGCTAGTATATAGTTTCACGATTTCCAGAAAGACACAGAATGCT 3221
1067 lGluArgLeuAsnThrSerLeuAlaPhePheLeuAsnAspLeuLeuServ 1084
|||||
3222 TGAAGACTCAATACAAAGCCTTGCAATTCTCTCAATGATCTGTGTCTG 3271
1084 alMetAspArgGlyPheValPheSerLeuIleLysSerCysTyrLysGln 1100
|||||
3272 TTATGCACAGAGGATTTGTTTATAGCCTTATAAAGTCTGCTATATAAG 3321
1101 ValSerSerLysLeuTyrSerLeuProAsnProSerValLeuValSerLe 1117
|||||
3322 GTGCTTCAAAAGCTTTACTCATTTACCGGAATCCAGTGTCTGTGTCCTT 3371
1117 uArgLeuAspPheLeuArgIleIleCysSerHisGluHisTyrValThrL 1134
|||||
3372 GAGCGTCGATTTTCTCGAATCATCTGCAGTCATGAGCACTATGTTACAT 3421
1134 euAsnLeuProCysSerLeuLeuThrProProAlaSerProSerProSer 1150
|||||
3422 TAAACTTACCCCTGACAGCTTACTTACTCCACCTGCATCTCCATCACCTTCT 3471
1151 ValSerSerAlaThrSerGlnSerSerGlyPheSerThrAsnValGlnAs 1167
|||||
3472 GTTCTCTCTGCAACATCTCAGAGTTCTGGAATTTTCTACGAATGTACAGA 3521

1167 pGlnLysIleAlaAsnMetPheGluLeuSerValProPheArgGlnGlnH 1184
|||||
3522 CCAAAAGATGGCAAAATATATGTTGAATTATCCGTGCGCTTCCGCCAACAGC 3571
1184 IStYrLeuAlaGlyLeuValLeuThrGluLeuAlaValIleLeuAspPro 1200
|||||
3572 ATTATTGGCAGCACTTGTTGTTAACAGAGCTGGCTGTCAATTTTAGACCCT 3621
1201 AspAlaGluGlyLeuPheGlyLeuHisLysLysValIleAsnMetValHi 1217
|||||
3622 GATGCTGAAGGACTGTTTGGATTGCATGAAGAAAGTCATCAATATGCTACA 3671
1217 sAsnLeuLeuSerSerHisAspSerAspProArgTyrSerAspProGlnI 1234
|||||
3672 CAATTACTCTCCAGTCACAGCACTCAGACCCGCGTACTCTGACCCCTCAGA 3721
1234 IeLysAlaArgValAlaMetLeuTyrLeuProLeuIleGlyIleIleMet 1250
|||||
3722 TAAAGGCTCGAGTGGCCATGTGTATCTACCTCTGATGGTATTCATCATG 3771
1251 GluThrValProGlnLeuTyrAspPheThrGluThrHisAsnGlnArgG 1267
|||||
3772 GAAACTGTACTCAGCTGTATGATTTTACAGAAACTCACAACTCAACAGAG 3821
1267 YArgProIleCysIleAlaThrAspAspTyrGluSerGluSerGlySerM 1284
|||||
3822 AAGACCAATTTGTATAGCCACTGATGATTAAGAAAGTGAGACGGAGTA 3871
1284 eTleSerGlnThrValAlaMetAlaIleAlaGlyThrSerValProGln 1300
|||||
3872 TGATTAAGCCAGACCGTTGGCATGGAATCGCAGGACATCGTCCCTCAA 3921
1301 LeuThrArgProGlySerPheLeuLeuThrSerThrSerGlyArgGlnHi 1317
|||||
3922 CTAAACAAGCCTGGCAGTTTCCCTCTCACGTCACAGAGTGGCAAGCAACA 3971
1317 sThrThrPheSerAlaGluSerSerArgSerLeuLeuIleCysLeuLeuT 1334
|||||
3972 CACTACCTTTTCAGCAGAATCAAGTCGAAGCCTTTTGATCTGTCTACTTT 4021
1334 rpValLeuLysAsnAlaAspGluThrValLeuGlnLysTrpPheThrAsp 1350
|||||
4022 GGGTCTCTCAAAATGCAGATGAACAGTTCTACAGAGTGTTTACAGAT 4071
1351 LeuSerValLeuGlnLeuAsnArgLeuLeuAspLeuLeuTyrLeuCysVa 1367
|||||
4072 CTCTCAGTCTTGCAGCTAAACCGGCTATTAGATCTGCTTTATCTCTGTGT 4121
1367 lSerCysPheGluTyrLysGlyLysLysValPheGluArgMetAsnSerL 1384
|||||
4122 GTCTTGCTTTGAGTATAAAGGGAAGAAAGTGTGGAACGAATGAATAGCT 4171
1384 euThrPheLysLysSerLysAspMetArgAlaLysLeuGlnGluAlaIle 1400
|||||
4172 TGACCTTTAAGAAATCAAAAGACATGAGAGCAAAAGCTTGAAAGAAGCTATT 4221
1401 LeuGlySerIleGlyAlaArgGlnGluMetValArgArgSerArgGlyG 1417
|||||
4222 CTGGAGACATAGGTGCCAGCAAGAAATGTATACGCCGAAGCCGAGACA 4271
1417 nLeuGluArgSerProSerGlySerAlaPheGlySerGlnGluAsnLeuA 1434
|||||
4272 GCTCGAGAGAACCCCATCTGCAAGTGCCTTTGGAAAGTCAGAAAAATTTGA 4321
1434 rGTrpArgLysAspMetThrHisTrpArgGlnAsnThrGluLysLeuAsp 1450
|||||
4322 GGTGAGAGAAAGATATGACTCACTGCGCTCAAAAACACAGAGAAAGCTTGAC 4371
1451 LysSerArgAlaGluIleGluHisGluAlaLeuIleAspGlyAsnLeuAl 1467
|||||
4372 AAATCAAGAGCAGAGATTGAACAGCAAGACACTGATGATGAGAAACCTTGGC 4421

1467 aThrGluAlaAsnLeuIleIleLeuAspThrLeuGluIleValValGlnT 1484
|||||
4422 TACAGAGCAAAACCTAAATCATTTTAGATACATTAGAGATTGTTGCAGA 4471
1484 hrValSerValThrGluSerLysGluSerIleLeuGlyGlyValLeuLys 1500
|||||
4472 CCGTTTCTGTAAAGGAATCCAAAGAGACATCTCTGTGGAGTGTCAAAA 4521
1501 ValLeuLeuHisSerMetAlaCysAsnGlnSerAlaValTyrLeuGlnHi 1517
|||||
4522 GTGCTACTACACAGCATGGCCCTGTAAACCAAGTCAGATTATCTACACA 4571
1517 sCysPheAlaThrGlnArgAlaLeuValSerLysPheProGluLeuLeuP 1534
|||||
4572 CTGTTTGTCTACAGAGAGAGCCTTGCTTCAAAAGTTTCTGAACCTCTTAT 4621
1534 heGluGluGluThrGluGlnCysAlaAspLeuCysLeuArgLeuLeuArg 1550
|||||
4622 TTGAAGAGAGACAGAGCAGTGTGATTTATGCTCAGGCTTCTCCGA 4671
1551 HisCysSerSerSerIleGlyThrIleArgSerHisProSerAlaSerle 1567
|||||
4672 CACTGTAGCAGTACATCGGTACAAATACGGTCAACACCCAGTGCCTCCCT 4721
1567 uTyrLeuLeuMetArgGlnAsnPheGluIleGlyAsnAsnPheAlaArgV 1584
|||||
4722 TTACCTACTAATGAGCAAAACTTTGAGATTGGGAATTAACTTTGGCAGGG 4771
1584 aLlyMetGlnValProMetSerLeuSerSerLeuValGlyThrSerGln 1600
|||||
4772 TTAAATGCAGGTACCAATGTCACTATCTCTTGTTGGGACATCTCAG 4821
1601 AsnPheAsnGluGluPheLeuArgArgSerLeuLysThrIleLeuThrTy 1617
|||||
4822 AATTTTAAAGAAATTTCTTAAGACGTTCTTAAGACTATATATGACATA 4871
1617 rAlaGluGluAspLeuGluLeuArgGluThrThrPheProAspGlnValG 1634
|||||
4872 TGCTGAAGAAAGATCTGGAATTGAGGGAACAACATTTCTGTGATCAGGTCC 4921
1634 lAspLeuValPheAsnLeuHisMetIleLeuSerAspThrValLysMet 1650
|||||
4922 AGCATCTGTTTCAATCTCCATATGATTCCTTCTGATACTGTGAATAATG 4971
1651 LysGluHisGlnGluAspProGluMetLeuIleAspLeuMetTyrArgIl 1667
|||||
4972 AAGGAACACCAGAGGATCTCCTGAATGTGATTGATCTAATGTACAGAAT 5021
1667 eAlaLysGlyTyrGlnThrSerProGluArgLeuThrTrpLeuGlnAsnM 1684
|||||
5022 TGCCAAGGGTTACCAAGACCTCTCCAGAGCGATTGACCTGGTTGCAGACA 5071
1684 eTAlaGlyLysHisSerGluArgSerAsnHisAlaGluAlaIleGlnCys 1700
|||||
5072 TGGCAGGCAAGCACTCAGAAGCAATCATGTCTGAAGCTGCACAGTGT 5121
1701 LeuValHisSerAlaAlaLeuValAlaGluTyrLeuSerMetLeuGluAs 1717
|||||
5122 CTAGTCCACTCAGCAGCACTGTGTGTAATATTGAGCATGTGGAAGA 5171
1717 pArgLysTyrLeuProValGlyCysValThrPheGlnAsnIleSerSera 1734
|||||
5172 CCGGAATATATCTCCTGTGGATGTGTAAACATTTCAGAATATATTCATCTA 5221
1734 snValLeuGluGluSerAlaValSerAspValValSerProAspGlu 1750
|||||
5222 ATGTTTTAGAAGAAATCTGCGGTCTCAGATGATGTGTATCTCCAGATGAA 5271
1751 GluGlyIleCysSerGlyLysTyrPheThrGluSerGlyLeuValGlyLe 1767
|||||
5272 GAAGGTATCTGCTCTGGAATAATACTTTACTGAGTCAGGACTTGTGGATT 5321
1767 uLeuGluGlnAlaAlaAlaSerPheSerMetAlaGlyMetTyrGluAlaIav 1784

|||||
5322 ACTGGAACAAGCAGCTGCTCCTCTCTATGCGTCGTCATGTGAAGCAG 5371
1784 alaAsnGluValTyrLysValLeuIleProIleHisGluAlaAsnArgAsp 1800
5372 TTAATGAAGTTTACAAGTACTTATTCCTATTCATGAAGCTAATCGGAT 5421
1801 AlaLysLysLeuSerThrIleHisGlyLysLeuGlnGluAlaPheSerIly 1817
5422 GCAAGAAGAACTATCCACAAATTCATGTAAACTTCAGAAGCATTCAGCAA 5471
1817 sIleValHisGlnSerThrGlyTrpGluArgMetPheGlyThrTyrPhea 1834
5472 AATTGTTCAATCAGAGTACTGGCTGGAGCGGATGTTGGCACCTATTTC 5521
1834 rGValGlyPheTyrGlyThrLysPheGlyAspLeuAspGluGlnGluPhe 1850
5522 GTGTGGCTTTTATGAACCAAGTTCGGGATTTGGATGAACAAGAATT 5571
1851 ValTyrLysGluProAlaIleThrLysLeuAlaGluIleSerHisArgIle 1867
5572 GTTACAAAGAGCCTGCATTAACCAAACTGCAGAGATATCTCACAGATT 5621
1867 uGluGlyPheTyrGlyGluArgPheGlyGluAspValValGluValIle 1884
5622 GGAGGATTTTACGGAGAAAGATTTGGAGAGGATGTGTTGAAGTAATCA 5671
1884 ysAspSerAsnProValAlaSpLysCysLysLeuAspProAsnLysAlaTyr 1900
5672 AAGACTCTAATCCTGTAGACAAAGTGTAATTAGATCCTAACAAGCATAT 5721
1901 IleGlnIleThrTyrValGluProTyrPheAspThrTyrGluMetLysAs 1917
5722 ATTCAGATTACCTATGTGAGGCATACTTTGACACATATGAGATGAAGGA 5771
1917 PaRgIleThrTyrPheAspLysAsnTyrAsnLeuArgArgPheMetTyrC 1934
5772 CAGAAACACCTATTTGCAGAAAAATTACAATCTTCGTCGATTCATGTA 5821
1934 ysThrProPheThrLeuAspGlyArgAlaHisGlyGluLeuHisGluGln 1950
5822 GTACACCCCTTACTTTAGATGGCCGTGCCCATGGGGAACCTTCATGAACA 5871
1951 PheLysArgLysThrIleLeuThrThrSerHisAlaPheProTyrIleIly 1967
5872 TTCAAAAGGAAGACCATCTCTGACTACGTCATGCGCTTCCTTATATTAA 5921
1967 sThrArgValAsnValThrHisLysGluGluIleIleLeuThrProIleG 1984
5922 AACAAAGGTCAAATGTCACTCATTAAGAAGATCATCTTAACCACAATTG 5971
1984 luValAlaIleGluAspMetGlnLysLysThrGlnGluLeuAlaPheAla 2000
5972 AAGTTGCTATTGAGGACATGCAGAAAAAGACACAGAGTTGCGATTGGCA 6021
2001 ThrHisGlnAspProAlaAspProLysMetLeuGlnMetValLeuGlnI 2017
6022 ACACATCAGGATCCCGCAGACCCCAAAATGCTTCAGATGGTACTCCAGGG 6071
2017 ySerValGlyThrThrValAsnGlnGlyProLeuGluValAlaGlnValP 2034
6072 ATCTGTAGGCACACAGTGAATCAGGGCCCTTTGGAAGTTGCCCAGGTTT 6121
2034 heLeuSerGluIleProSerAspProLysLeuPheArgHisAsnLys 2050
6122 TTCTGTCTGAAATACCTAGTGACCCCAAAGCTCTTCAGACATCATTAATAA 6171
2051 leuArgLeuCysPheLysAspPheThrLysArgCysGluAspAlaLeuAr 2067
6172 CTGCGACTCTGCTTTAAAGATTTTACTAAAGGTGTGAAGATGCTTTAAG 6221
2067 gLysAsnLysSerLeuIleGlyProValGlnLysGluTyrGlnArgGluL 2084
|||||

6222 AAAAATAAGAGCTTAATTGGCCCGTTCAAAAGAGTATCAAGGGAAT 6271
2084 euGlyLysLeuSerSerPro 2090
6272 TGGGAAACTATCTTCGCCT 6291
seq_name: /SIDs1/gcgdata/hold-geneseq/geneseqn-emb1/NA2001A.DAT:AAH43851
seq_documentation_block:
ID AAH43851 standard; cDNA; 6372 BP.
XX
AC AAH43851;
XX
DT 04-SEP-2001 (first entry)
XX
DE Human CLASP-7 encoding cDNA sequence SEQ ID NO:1.
XX
KW Human; CLASP-7; cadherin-like asymmetry protein; immune response;
KW neuroprotective; antidiabetic; immunosuppressive; antirheumatic;
KW antiarthritic; hypotensive; anti-HIV; cytostatic; immunostimulant;
KW antianaemic; antiinflammatory; ophthalmological; nephrotrophic;
KW antithyroid; antiasthmatic; antiallergic; antibacterial; gene therapy;
KW chromosome 19q13.2; autoimmune disease; multiple sclerosis; toxemia;
KW juvenile diabetes; rheumatoid arthritis; pruritic urticarial papule;
KW hypertension; Rh incompatibility; ss.
XX
OS Homo sapiens.
XX
FH Key location/Qualifiers
FT CDS 13..6156
FT /*tag= a
FT /*product= "CLASP-7"
XX
PN WO200142295-A2.
XX
PD 14-JUN-2001.
XX
PF 13-DEC-2000; 2000WO-US34152.
XX
PR 13-DEC-1999; 99US-0170453.
PR 14-JAN-2000; 2000US-0176195.
PR 14-FEB-2000; 2000US-0182296.
PR 11-APR-2000; 2000US-0196267.
PR 11-APR-2000; 2000US-0196460.
PR 11-APR-2000; 2000US-0196527.
PR 11-APR-2000; 2000US-0196528.
PR 11-APR-2000; 2000US-0547276.
PR 13-OCT-2000; 2000US-0240503.
PR 13-OCT-2000; 2000US-0240508.
PR 13-OCT-2000; 2000US-0240539.
PR 13-OCT-2000; 2000US-0240543.
XX
PA (ARBO-) ARBOR VITA CORP.
XX
PI Lu P, Garman JD, Candia AF;
XX
DR WPI; 2001-381641/40.
DR P-PSDB; AAB99541.
XX
PT Novel cadherin-like asymmetry protein-7 and polynucleotides encoding
PT the polypeptide, useful for treating autoimmune disease,
PT hypersensitivity, preventing transplant rejection by modulating immune
PT response -
XX
PS Claim 1; Fig 5; 151pp; English.
XX
CC The present invention describes a human cadherin-like asymmetry protein
CC (CLASP), designated CLASP-7. The CLASP-7 protein (I) and its encoding
CC nucleotide sequence (II) have activities including: neuroprotective;
CC antidiabetic; immunosuppressive; antirheumatic; antiarthritic; anti-HIV;
CC hypotensive; cytostatic; immunostimulant; antianaemic; antiinflammatory;
CC ophthalmological; nephrotrophic; antithyroid; antiasthmatic;
CC antiallergic; and antibacterial. (II) and CLASP-7 antibodies (III) are

1427 TCTTCAGTTCCTGGCTGACATGAGCGCCGCTGCTCCCTGCTGGCGGA 1476
502 LeuArgProIleThrAlaGlnLeuLysIleAspIleSerProAlaProG1 518
1477 CTACGTCCTGTGACTGCCAGCTCAAGATGCACATTCTCCGGCTCTGA 1526
518 uAsnProHisTyrCysLeuThrProGluLeuGlnValLysLeuTyrP 535
1527 AAATCCCACACTTCGGCTCTCCCTGAGCTGCTTCATATCAAGCCCTACC 1576
535 roAspSerArgValArgProThrArgGluIleLeuGluPheProAlaArg 551
1577 CGGACCCCGAGGGCGGCCCAAGAGAGATTCTGGAGTTCCCGCCCGCC 1626
552 AspValTyrValProAsnThrTyrTyrArgAsnLeuLeuTyrIleTyrPr 568
1627 GAAGTCTATGCCCCCAACCACTACAGGAACCTGCTGTACGTGTACCC 1676
568 ogInserLeuAsnPheAlaAsnArgGlnGlySerAlaArgAsnIleThrV 585
1677 GCACAGCCTCAACTTCAGCAGCCGCCAGGGCTCCGTGCGCAACCTTGCTG 1726
585 alLysValGlnPheMetTyrGlyGluAspProSerAsnAlaMetProVal 601
1727 TCGGAGTGCAGTACATGACAGCGGAGGACCCAGCCAGGCTCTGCCGCTC 1776
602 IlePheGlyLysSerSerCysSerGluPheSerLysGluAlaTyrThrAl 618
1777 ATCTTTGGCAAGTCCAGCTGCAGTGAATTACCCCGGAGGCCCTTCACACC 1826
618 aValValTyrHisAsnArgSerProAspPheHisGluGluIleLysValL 635
1827 GGTGGTCTACCATACAAAGTCCCGGAGTTCTACGAGGAGTTCAAGCTGC 1876
635 ySLeuProAlaThrLeuThrAspHisHisIleLeuPheThrPheTyr 651
1877 ATCTTCCAGCCTGCGTGACAGAGAACCATCACCTGCTGTTCACCTTCTAC 1926
652 HisValSerCysGlnGlnLysGlnAsnThrProLeuGluThrProValG1 668
1927 CATGTCAGCTGCCAGCCCCGGCGGCACTGCCCTGGAGACACCCGTGGG 1976
668 yTyrThrTrrPilePrometLeuGlnAsnGlyArgLeuLysThrGlyGlnP 685
1977 CTTTACTTGATCCCACTGCTGCAGCACGGCGCCCTGAGGACCGGCCCT 2026
685 heCysLeuProValSerLeuGluLysProProGlnAlaTyrSerValLeu 701
2027 TCTGTCTCCAGTGTGTGGAGCCAGCCGCCGCCCACTATTCCGTGCTC 2076
702 SerProGluValProLeuProGlyMetLysTrrPValAspAsnHisLysG1 718
2077 ACACCCGATGTGGCGTTCCGGGCATGCGCTGGTGAGCGGTCAACAAGGG 2126
718 yValPheAsnValGluValAlaValAspSerIleHisThrGlnAspP 735
2127 CGTGTTCAGTGTGAGCTCACAGCCGTGCTCTGTGCACCCCGCAGGACC 2176
735 roTyrLeuAspLysPhePheAlaLeuValAsnAlaLeuAspGluHisLeu 751
2177 CCTACTGGACAATTTCTTCACCTGGTGACACGTCTGAGAGGAGGCC 2226
752 pheProValArgIleGlyAspMetArgIleMetGluAsnAsnLeuGluAs 768
2227 TTCCCATTTCCGGCTCAAGGACACTGTGCTGAGCGAGGGCAACGTGAGCA 2276
768 ngLueLysSerSerIleSerAlaLeuAsnSerSerGlnLeuGluProv 785
2277 GGAGCTGGGGCCAGTCTTGCAGCACTGCGCCTGGCGACCCCGAAACCC 2326
785 alValArgPheLeuHisLeuLeuAspLysLeuIleLeuLeuValIle 801

2327 TTGTGGCCTTCTCCACACAGCTGCTGGACAAGCTCTGCGCTCTGTCATC 2376
802 ArgProProValIleAlaGlyGlnIleValAsnLeuGlyGlnAlaSerPh 818
2377 AGGCCCCGATCATCAGTGGCCACAGATTGTGAACCTGGGCCGTGAGCCTT 2426
818 eGluAlaMetAlaSerIleIleAsnArgLeuHisLysAsnLeuGluGlyA 835
2427 TGAAGCAATGGCCCATGTAGTCAGCCTGTGTTCACCGGAGCCTGAGGCGAG 2476
835 snHisAspGlnHisGlyArgAsnSerLeuLeuAlaSerTyrIleHisTyr 851
2477 CCCAGGATGCCCGCGGTCACTGCCACACAGCTGGCTGGCTTACGTCACCTAC 2526
852 ValPheArgLeuProAsnThrTyrProAsnSerSerSerProGlyProG1 868
2527 GCCTTTCGCTTCTCTGGACACTGAGCCACGCTCCCGGATGGGCCCT.. 2574
868 yGlyLeuGlyGlySerValHisTyrAlaThrMetAlaArgSerAlaValA 885
2575CCAGTGACAGTGCAGGCTGCCACACTGGCCCGTGGCTCTGGTC 2617
885 rgProAlaSerLeuAsnLeuAsnArgSerArgSerLeuSerAsnSerAsn 901
2618 GCCCCGCAAGCCTTACTCTGCGCGCTTCCAAGAGCATCAGCAGCAAC 2667
902 ProAspIleSerGlyThrProThrSerProAspAspGluValArgSerIl 918
2668 CCTGACCTCGCCGTGGCCCTGGCTCTGTGATGACGAGGTTTCCCGCAT 2717
918 eiLeGlySerLysGlyLeuAspArgSerAsnSerTrrPValAsnThrGlyG 935
2718 CCTGGCCAGC..... 2727
935 lyProLysAlaAlaProTrrPlySerAsnProSerProSerAlaGluSer 951
2727 2727
952 ThrGlnAlaMetAspArgSerCysAsnArgMetSerSerHisThrGluTh 968
2727 2727
968 rSerSerPheLeuGlnThrLeuThrGlyArgLeuProThrLysLysLeuP 985
2728AAGCTGC 2734
985 heHisGluGluLeuAlaLeuGlnTrrPValAlCysSerGlySerValArg 1001
2735 TTCACGAGAGAGCTGGCTCTGCAGTGGGTGTCAGCAGCAGTGGCCGTACGC 2784
1002 GluSerAlaLeuGlnGlnAlaTrrPhePhePheGluLeuMetValLysSe 1018
2785 GAGGCCATCCTCCAGCAGCAGCCTGTTCTTCTTCCAGCTCATGTGTAAGAG 2834
1018 rMetValHisHisLeuTyrPheAsnAspLysLeuGluAlaProArgLysS 1035
2835 TATGGCGCTGCACCTGCTGTTGGCCAGGACTAGACACACCCCGCAAGC 2884
1035 eArgPheProGluArgPheMetAspAspIleAlaAlaLeuValSerThr 1051
2885 TGCCTTCCCGGACGCTTCTCTGAGACATCACTGCTTGGTGGGCTCT 2934
1052 IleAlaSerAspIleValSerArgPheGlnLysAspThrGluMetValG1 1068
2935 GTGGCGCTGAGGTATCACCCGTGTCCACAAAGATGTGAGACTGGCCGA 2984
1068 uArgLeuAsnThrSerLeuAlaPhePheLeuAsnAspLeuLeuSerValM 1085
2985 GCACCTCAACGCGCAGCCTGGCTTCTTCTCAGTGACCTTCTGCCCTGG 3034
1085 etAspArgGlyPheValPheSerLeuIleLysSerCysTyrLysGlnVal 1101
3035 TGCACGGGGCTTTGTCTTACGCCCTGGTCCGGCCCACTACAGCAGGTG 3084

1102	SerSerLysLeuTyrSerLeuProAsnProSerValLeuValSerLeuAr	1118
3085	GCACGCGGCTCCAGTCGTCCCTCATCCAGCAGCCCTGTGACCTGCG	3134
1118	GleuAspPheLeuArgIleIleCysSerHisGluHisTyrValThrLeuA	1135
3135	CATGGAATTCAACCGCATCTGTGCAGCCAGCAGACTAGTACCTCA	3184
1135	snLeuProCysSerLeuLeuThrProProAlaSerProSerProSerVal	1151
3185	ACCTCCCTGTGTCGCCCTGTCACTCCAGCCTCGCCCTCCCTCTGTG	3234
1152	SerSerAlaThrSerGlnSerSerClyPheSerThrAsnValGlnAspGl	1168
3235	TCCCTCCACCACCTCCAGAGAGCTTCACCTTCTCCAGCCAAAGCCCGGACC	3284
1168	nLysIleAlaAsnMetPheGluLeuSerValProPheArgGlnHisT	1185
3285	CAAGCTGACCACGATGTTGCACTAGTGGACCATTCGGCGACGACGACT	3334
1185	yrLeuAlaGlyLeuValLeuThrGluLeuAlaValIleLeuAspProAsp	1201
3335	TCCTAGCTGGGCTCCTGTGTCAGGAGCTGGCAGCTGCCCTGCAACTGAG	3384
1202	AlaGluGlyLeuPheGlyLeuHisLysLysValIleAsnMetValHisAs	1218
3385	GCTGAAGGGGCAATTCCTGTGTGCACAAGAAGGCATCAGTGTGTGCACAG	3434
1218	nLeuLeuSerSerHisAspSerAspProArgTyrSerAspProGlnIleL	1235
3435	CCTGCTATGTGGCCATGACACTGACACCCCGCTACGCCGAGGCCACTGTGA	3484
1235	ysAlaArgValAlaMetLeuTyrLeuProLeuIleGlyIleIleMetGlu	1251
3485	AGGCTCGTGTGGCCGAGCTGTACCTGCCACTGCTTCGATGTGCACGGGAT	3534
1252	ThrValProGlnLeuTyrAspPheThrGluThrHisAsnGlnArgGlyAr	1268
3535	ACCTTGCACAGGCTGCATGACTTGTCTGAGGGCCAGGTCAGCGGTCAAG	3584
1268	gProIleCysIleAlaThrAspAspTyrGluSerGluSer.....Glys	1283
3585	ACTGGCCTCATATGCTTGACTTCAGACACAGAGAAGGCGAAGGGACATTGGCG	3634
1283	erMetIleSerGlnThrValAlaMetAlaIleAlaGlyThrSerValPro	1299
3635	GTAACCATCAACCCCTCTGTGGCCATGGCCATTGCTGTGGCCCCCTTA..	3681
1300	GlnLeuThrArgProGlySerPheLeuLeuThrSer.....Th	1312
3682GCCCTGGCTCCCGGGCCAGCATCTCCAGGGGCCACCAAC	3722
1312	rSerGlyArgGlnHisThrThrPheSerAlaGluSerSerArgSerLeuL	1329
3723	GCGTCTCTCGCGCAGGCTGTGCCCTCTCTGTGAGTCAAGCCGACCTTGC	3772
1329	eulIleCysLeuLeuTyrValLeuLysAsnAlaAspGluThrValLeuGln	1345
3773	TGGCGTGTGTGTGTGGGTCTGAAAAAACACGACCGCGGCTCTCTGCAG	3822
1346	LysTrpPheThrAspLeuSerValLeuGlnLeuAsnArgLeuLeuAspLe	1362
3823	CGCTGGGCGCACTGACCTGACACATCCCCCAGCTGGGACGTCGTGGACTT	3872
1362	uLeuTyrLeuCysValSerCysPheGluTyrTyrLysGlyLysLysValPheG	1379
3873	GCTGTACCTTTGCCCTGGCTGCTTTGAGTACAAGGGGAAAAAGGCCCTTGG	3922
1379	luArgMetAsnSerLeuThrPheLysLysSerLysAspMetArgAlaLys	1395
3923	AACGCATCAACAGCTCATATTCAAAAAATCTGTGATATGAAGCGCGG	3972

1396	LeuGIuGluaIaIaIeLeuGIySerIleGIyAlaArgGIuGIuMetValAr	1412
1397		1413
1398		1414
1399		1415
1400		1416
1401		1417
1402		1418
1403		1419
1404		1420
1405		1421
1406		1422
1407		1423
1408		1424
1409		1425
1410		1426
1411		1427
1412	gArSerArGIyGIuLeuGIuArgSerProSerGIySerAlaPheGIys	1429
1413		1430
1414		1431
1415		1432
1416		1433
1417		1434
1418		1435
1419		1436
1420		1437
1421		1438
1422		1439
1423		1440
1424		1441
1425		1442
1426		1443
1427		1444
1428		1445
1429	erGIuGIuAsnLeuArgTrpArgLysAspMetThrHisTrpArgGIuAsn	1445
1430	:::	1446
1431	:::	1447
1432	:::	1448
1433	:::	1449
1434	:::	1450
1435	:::	1451
1436	:::	1452
1437	:::	1453
1438	:::	1454
1439	:::	1455
1440	:::	1456
1441	:::	1457
1442	:::	1458
1443	:::	1459
1444	:::	1460
1445	:::	1461
1446	ThrGIuLysLeuAspLysSerArgAlaGIuIleGIuHisGIuAlaLeuI	1462
1447	:::	1463
1448	:::	1464
1449	:::	1465
1450	:::	1466
1451	:::	1467
1452	:::	1468
1453	:::	1469
1454	:::	1470
1455	:::	1471
1456	eAspGIyAsnLeuAlaThrGIuAlaAsnLeuIleIleLeuAspThrLeuG	1479
1457	:::	1480
1458	:::	1481
1459	:::	1482
1460	:::	1483
1461	:::	1484
1462	:::	1485
1463	:::	1486
1464	:::	1487
1465	:::	1488
1466	:::	1489
1467	:::	1490
1468	:::	1491
1469	:::	1492
1470	:::	1493
1471	:::	1494
1472	:::	1495
1473	:::	1496
1474	:::	1497
1475	:::	1498
1476	:::	1499
1477	:::	1500
1478	:::	1501
1479	:::	1502
1480	:::	1503
1481	:::	1504
1482	:::	1505
1483	:::	1506
1484	:::	1507
1485	:::	1508
1486	:::	1509
1487	:::	1510
1488	:::	1511
1489	:::	1512
1490	:::	1513
1491	:::	1514
1492	:::	1515
1493	:::	1516
1494	:::	1517
1495	:::	1518
1496	:::	1519
1497	:::	1520
1498	:::	1521
1499	:::	1522
1500	:::	1523
1501	:::	1524
1502	:::	1525
1503	:::	1526
1504	:::	1527
1505	:::	1528
1506	:::	1529
1507	:::	1530
1508	:::	1531
1509	:::	1532
1510	:::	1533
1511	:::	1534
1512	:::	1535
1513	:::	1536
1514	:::	1537
1515	:::	1538
1516	:::	1539
1517	:::	1540
1518	:::	1541
1519	:::	1542
1520	:::	1543
1521	:::	1544
1522	:::	1545
1523	:::	1546
1524	:::	1547
1525	:::	1548
1526	:::	1549
1527	:::	1550
1528	:::	1551
152		

|||||
4852 GCCGAGGCCGCCAGTGCATGTTGCACGCGGCCCTCGTGCTGAGTA 4901
1711 rleuSerMetleuGluAspArgLysTyrleuProValGlyCysValThrp 1728
|||||
4902 CCTCGCCCTGTGAGGAGCAGCGCCACTGCGGTGGCTGCGTTCCCT 4951
1728 heGlnAsnIleSerSerAsnValleuGluSerAlaValSeraSpasp 1744
|||||
4952 TTCAGACATCTCATCCAACGTCGTAGAGGAGTCCGCCATCTCCGACGAC 5001
1745 ValValSerProAspGluGluGlyIleCysSerGlyLysTyrPheThrgl 1761
:::|||||
5002 ATCCTGTGCGCCGACGAGGAGGCGCTTCTCTCCGGGAAGCACTTCACTGA 5051
1761 uSerGlyLeuValGlyLeuLeuGlnAlaAlaSerPheSerMet 1778
| |||||
5052 GCTGGGCTGTAGGCTTGTGTGAACAGGCGCGCTACTTCAACCATGG 5101
1778 laGlyMetYrGluAlaValAsnGluValTyrLysValleuIleProIle 1794
:::|||||
5102 GCGGCTCTACGAGCGCGTGAATGAGTCTACAAGACCTCATCCCCATC 5151
1795 HisGluAlaAsnArgAspAlaLysLysLeuSerThrIleHisGlyLysLe 1811
|||||
5152 CTGGAAGCCCAACCGTGACTACAGAAGCTGCGCGGTGCACGCAAACT 5201
1811 uGlnGluAlaPheSerLysIleValHisGlnSerThrGlyTrpGluArgm 1828
|||||
5202 GCAGGAGGCTTCAACAGATCATGCACAGAGTTCCGGCTGGGAGCGCG 5251
1828 etPheGlyThrTyrPheArgValGlyPheTyrGlyThrLysPheGlyasp 1844
:::|||||
5252 TGTTCGGGCGTATTTCCGCGTGGCTTCTACGCGGCCCACTTCGCTGAC 5301
1845 leuAspGluGlnGluPheValTyrLysGluProAlaIleThrLysleuAl 1861
|||||
5302 CTGATGAGCAGGAGTTTGTGTACAGAGCATCATGCACGAAGCTGGC 5351
1861 agluIleSerHisArgleuGluGlyPheTyrGlyLysArgPheGlyGlu 1878
|||||
5352 AGAGATCTCACACCGGCTGGAGAGTTCTACACGGAGAGATTGGCGACG 5401
1878 spValValGluValIleLysAspSerAsnProValAspLysCysLysleu 1894
|||||
5402 ACGTCGTTGAGATTATCAAGACTCTTACCCTGTGACAAAGTCCAAGCTT 5451
1895 AspProAsnLysAlaTyrIleGlnIleThrTyrValGluProTyrPheAs 1911
||| :::|||||
5452 GACTCACAAAGGCTACATCCAGATCAGTATGTGAACCGTACTTTGA 5501
1911 pThrTyrGluMetLysAspArgIleThrTyrPheAspLysAsnTyrAsnL 1928
|||||
5502 TACCTACGAGCTCAAGGAGCGGCTGACTTGTGACCGCACTATGGGC 5551
1928 euArgArgPheMetTyrCysThrProPheThrleuAspGlyArgAlaHis 1944
|||||
5552 TTCGCACATTCCTGTTCTGCACGCGCTTACGCGCGGATGGCGGCACAC 5601
1945 GlyGluLeuHisGluGlnPheLysArgLysThrIleleuThrThrSerH 1961
|||||
5602 GGGGAGCTGCCCGAGCAACACAGCGTTAAGACGCTGCTCAGCACCGACCA 5651
1961 sAlaPheProTyrIleLysThrArgValAsnValThrHisLysGluGluI 1978
|||||
5652 CGCCTTCCCTACATCAAGACTCGCATCCGTGTGTGCACCGGAGGAGA 5701
1978 leIleLeuThrProIleGluValAlaIleGluAspMetGlnLysLysThr 1994
:::|||||
5702 CGGTGCTGAGCCAGTGAGGTGCGCATCGAGACATGCAGAGAAAGACA 5751
1995 GlnGluLeuAlaPheAlaThrHisGlnAspProAlaAspProLysMetle 2011
:::|||||

5752 CGGAGCTGGCCTTTGCCACCGAGCAGGACCCACAGATGCTAAGATGCT 5801
2011 uGlnMetValleuGlnGlySerValGlyThrThrValAsnGlnGlyProL 2028
|||||
5802 ACAGATGGTGTTCAGGCGCTCTGTAGGGCCACCGTGAACAGGCTCCC 5851
2028 euGluValAlaGlnValPheLeuSerGluIleProSerAspProLysleu 2044
|||||
5852 TGGAGGTGGCCAGGCTTTTTAGCAGAGATCCCGGAAGACCCCAAGCTC 5901
2045 PheArgHisHisAsnLysleuArgLeuCysPheLysAspPheThrLysAr 2061
|||||
5902 TTCGGGCATCACACAATTCGCGCTCTGCTTCAAGACTTCTGCAGAA 5951
2061 gCysGluAspAlaLeuArgLysAsnLysSerleuIleGlyProValGlnL 2078
: |||||
5952 ATGTGAGGATGCGCTGCGGAAAAATAAGGCGCTGATTGGCGCGGACCA 6001
2078 ysgLutYrGlnArgGluLeu 2084
|||||
6002 AGGAGTACCACCGTGAGCTG 6021
seq_name: /SIDSl/gcgdata/hold-geneseq/geneseqn-emb1/NA2001A.DAT:AAH41953
seq_documentation_block:
ID AAH41953 standard; DNA; 4200 BP.
XX AAH41953;
AC 30-AUG-2001 (first entry)
DT 30-AUG-2001 (first entry)
XX CLASP-3 nucleotide fragment.
DE Human; cadherin-like asymmetry protein; CLASP; CLASP-3; immune response;
XX cell surface molecule; transmembrane protein; immunosuppressive; vaccine;
KW antiinflammatory; antiarthritic; antianaemic; dermatological; uropathic;
KW ophthalmological; antirheumatic; nephrotropic; antidiabetic;
KW neuroprotective; antiasthmatic; antibacterial; antisense therapy;
XX gene therapy; chromosome 1p31.1; ds.
XX Unidentified.
OS WO200142297-A2.
PN 14-JUN-2001.
XX 13-DEC-2000; 2000WO-US34171.
PF 13-DEC-1999; 99US-0170453.
PR 14-JAN-2000; 2000US-0176195.
PR 14-FEB-2000; 2000US-0182296.
PR 11-APR-2000; 2000US-0196267.
PR 11-APR-2000; 2000US-0196460.
PR 11-APR-2000; 2000US-0196527.
PR 11-APR-2000; 2000US-0196528.
PR 11-APR-2000; 2000US-0547276.
PR 13-OCT-2000; 2000US-0240503.
PR 13-OCT-2000; 2000US-0240508.
XX (ARBO-) ARBOR VITA CORP.
PA Lu P, Garman JD, Candia AF;
XX WPI, 2001-375003/39.
XX Novel isolated cadherin-like asymmetry protein (CLASP) -3, useful for
PT useful in preventing or treating a CLASP-3-mediated disease e.g.
PT autoimmune disease
XX
XX Disclosure; Fig 9C; 189pp; English.
XX The present invention describes an isolated polypeptide (I) comprising a
CC nucleotide (nt) sequence (Sla) that has at least 90% identity to the

CC sequence given in AAH41934 and is immunologically cross-reactive with
CC the derived amino acid (aa) sequence (Sib) given in AAB99495 or shares a
CC biological function with native CLASP-3. (I) has immunosuppressive,
CC antiinflammatory, antiarthritic, antianaemic, dermatological, uropathic,
CC ophthalmological, antirheumatic, nephrotropic, antithyroid, antidiabetic,
CC neuroprotective, antiasthmatic and antibacterial activities, and can be
CC used in antisense therapy, vaccine production and gene therapy. CLASP-3
CC related sequences can be used in preventing or treating a CLASP-3
CC mediated disease, preferably an autoimmune disease by inhibiting an
CC immune response. The autoimmune disease is caused or exacerbated by
CC increased activity of T helper 1 lymphocytes (TH1). AAH41893 to AAH41953
CC and AAB99491 to AAB99507 represent sequences which are used in the
CC exemplification of the present invention. CLASP-3 is localised in the
CC chromosome location 1p31.1.

XX
SQ Sequence 4200 BP; 1262 A; 844 C; 907 G; 1187 T; 0 other;

alignment_scores:

Quality: 6171.50 Length: 1226
Ratio: 5.088 Gaps: 2
Percent Similarity: 98.940 Percent Identity: 98.206

alignment_block:

US-09-737-246-2 x AAH41953 ..

Align seg 1/1 to: AAH41953 from: 1 to: 4200

866 GLYProGLYGLYLeuGLYGLYSerValHISYrAlaThrMetAlaArgse 882
|||||
3 GGGCCCCGTCGCCGTCGGTCGACGGTATCGATAGACTTGATATCGAATTC 52
882 rAlaValArgProAlaSerLeuAsnLeuAsnArgSerArgSerLeuSera 899
|||||
53 GGCA.....CGAGTAAATCGTTCTCGAAGCCTTAGTA 84
899 snSerAsnProAspIleSerGlyThrProThrSerProAspAspGluVal 915
|||||
85 ATAGCAATCCAGATATATGTGGACTCCACGTCACCAAGATGAAAGTT 134
916 ArgSerIleIleGlySerLysGlyLeuAspArgSerAsnSerTrpValAs 932
|||||
135 CGATCAATCATCGGAGCTAAGGGTTAGATCGCTCCAATTCCGTGGTTAA 184
932 nthrGLYGLYProLysAlaAlaProTrpGLYSerAsnProSerProSera 949
|||||
185 CACTGGTGGTCCAAAGCTGCCCATGGGATCCAAACCCAGTCCCAAGTG 234
949 IagIuSerThrGlnAlaMetaspArgSerCysAsnArgMetSerSerHis 965
|||||
235 CAGAATCAACACAGGCTATGGATCGAAGTTGTAATCGTATGCTTCGCAC 284
966 ThrGluThrSerSerPheLeuGlnThrLeuThrGlyArgLeuProThrLy 982
|||||
285 ACAGAGACGTCAGTTCCTTACAACACATTAAACGGGACGCTTACCACATA 334
982 sLysLeuPheHisGluGluLeuAlaLeuGlnTrpValValCysSerGlyS 999
|||||
335 AAAGCTTTTTCAGAGAGAGCTGGCTTTCAGTGGGTTGTTGCAGTGGCA 384
999 erValArgGluSerAlaLeuGlnGlnAlaTrpPhePhePheGluLeuMet 1015
|||||
435 GTAAAGAGCATGGTGCACATTATCTTAAATGATAAAGCTTGAGGCTTC 484
1016 ValLysSerMetValHisHisLeuTyrrPheAsnAspLysLeuGluAlaPr 1032
|||||
1032 oArgLysSerArgPheProGluArgPheMetaspIleAlaAlaLeuV 1049
|||||
485 AAGGAAAGTGTGTTTCCAGAAAGCTTTCATGAGATGACATTGCAGCTCTTG 534
1049 alSerThrIleAlaSerAspIleValSerArgPheGlnLysAspThrGlu 1065

|||||
535 TCAGCAGCATTTGCTAGTATAGTTTCCAGATTTTCAGAGACACAGAA 584
1066 MetValGluArgLeuAsnThrSerLeuAlaPhePheLeuAsnAspLeu 1082
|||||
585 ATGGTTGAGAGACTCAATACAGCCCTTGCAATCTTCTCATGATCTGTT 634
1082 uSerValMetaspArgGlyPheValPheSerLeuIleLysSerCysTrpL 1099
|||||
635 GTCTGTATGACAGAGAGATTTGTTTAGCCTTATAAGTCCGTGATATA 684
1099 ySGlnValSerSerLysLeuTyrrSerLeuProAsnProSerValLeuVal 1115
|||||
685 AACAGGTGCTTCAAAGCTTACTCATTTACCGAATCCAGTGTCTGTG 734
1116 SerLeuArgLeuAspPheLeuArgIleIleCysSerHisGluHisTyrrVa 1132
|||||
735 TCCTTGAGGCTGGATTTTCTACGAATCATCTGCAGTCAAGACACTATGT 784
1132 lThrLeuAsnLeuProCysSerLeuLeuThrProProAlaSerProSerP 1149
|||||
785 TACATTTAAACTTACCCTGCAGCTTACTTACTCCACCTGCATCTCCATCAC 834
1149 roSerValSerSerAlaThrSerGlnSerSerGlyPheSerThrAsnVal 1165
|||||
835 CTTCCTGTTCTCTGCAACATCTCAGAGTTCTGGATTTTCTACGAATGTA 884
1166 GlnAspGlnLysIleAlaAsnMetPheGluLeuSerValProPheArgG 1182
|||||
885 CAAGACCAAAAGATTGCAAAATATGTTGAATTATCCGTGCTTCCGCCA 934
1182 nGlnHisTyrrLeuAlaGlyLeuValLeuThrGluLeuAlaValIleLeuA 1199
|||||
935 ACAGCAATTATTTGGCAGAGACTGTGTTAAACAGAGCTGCTGCATTTTAG 984
1199 sPProaspAlaGluGlyLeuPheGlyLeuHisLysLysValIleAsnMet 1215
|||||
985 ACCCTGATGCTGAAGAGACTGTTGGATTGCAATAAGAACTCATCAATATG 1034
1216 ValHisAsnLeuLeuSerSerSerHisaspSerAspProArgTyrrSerAspPr 1232
|||||
1035 GTACACAATTACTCTCCAGTCACGACTCAGACCCCGCGTACTCTGACCC 1084
1232 oGlnIleLysAlaArgValAlaMetLeuTyrrLeuProLeuIleGlyIleI 1249
|||||
1085 TCAGATAAAGGCTCGAGTGGCATGTTGTATCTACCTCTGATTGGTATTA 1134
1249 lMetGluThrValProGlnLeuTyrrAspPheThrGluThrHisAsnGln 1265
|||||
1135 TCATGGAACCTGTACCTCAGCGTATGATTTTACAGAAACTCACAAATCAA 1184
1185 ArgGlyArgProIleCysIleAlaThrAspAspTyrrGluSerGluSerG 1282
|||||
1282 ySerMetIleSerGlnThrValAlaMetAlaIleAlaGlyThrSerValP 1299
|||||
1235 AAGTATGATTAAGCCAGACCGTTGCCATGGCAATCGCAGGACATCGGTCC 1284
1299 roGlnLeuThrArgProGlySerPheLeuLeuThrSerThrSerGlyArg 1315
|||||
1285 CTCACCTAACAAGCGCTGGCAGTTTCCCTCCACGTCACAGAGAGTGGCAGG 1334
1316 GlnHisThrThrPheSerAlaGluSerSerArgSerLeuLeuIleCysLe 1332
|||||
1335 CAACACACTACCTTTTCAGCAGAAATCAAGTCGAAGCCTTTTGATCTGTCT 1384
1332 uLeuTrpValLeuLysAsnAlaaspGluThrValLeuGlnLysTrpPheT 1349
|||||
1385 ACTTTGGTCTCTCAAAAATGCGATGAACAGAGTCTTACAGAAAGTGGTTTA 1434
1349 hrAspLeuSerValLeuGlnLeuAsnArgLeuLeuAspLeuLeuTyrrLeu 1365
|||||

1435 CAGATCTCTCAGTCTTGACGCTAAACGGCTATTAGATCTGCTTATCTC 1484
1366 CysValSerCysPheGluTyrLysGlyLysLysValPheGluArgMetAs 1382
1485 TGTGTGCTTGCTTTGAGTATAAAGGAAAAAGTGTTTGAAAGCAATGAA 1534
1382 nSerLeuThrPheLysLysSerLysAspMetArgAlaLysLeuGluAla 1399
1535 TAGCTTGACCTTTAAGAAATCAAAAGACATGAGAGCAAGCTTGAAAGAA 1584
1399 laIleLeuGlySerIleGlyAlaArgGlnGluMetValArgArgSerArg 1415
1585 CTATTTCTTGGAGCATAGTGTCCAGCAAGAAATGTACGCCGAAGCCGA 1634
1416 GlyGlnLeuGluArgSerProSerGlySerAlaPheGlySerGlnGluAs 1432
1635 GGACAGCTCGAGAGAGCCCATCTGGAAGTGCCTTGGAAGTCAAGAAAA 1684
1432 nLeuArgTrpArgLysAspMetThrHisTrpArgGlnAsnThrGluLysL 1449
1685 TTTGAGGTGGAGGAAGATATGACTCAGTGGCGTCAAAACACAGAGAAGC 1734
1449 euAspLysSerArgAlaGluIleGluHisGluAlaLeuIleAspGlyAsn 1465
1735 TTGACAAATCAAGAGCAGAGATTTGAACACGAGCACTGATTGATGGAAC 1784
1466 LeuAlaThrGluAlaAsnLeuIleIleLeuAspThrLeuGluIleValVa 1482
1785 CTGGCTACAGAAAGCAACCTTAATCATTTAGATACATTAGAGATTGTGT 1834
1482 IGlnThrValSerValThrGluSerLysGluSerIleLeuGlyGlyValL 1499
1835 TCAGACCGTTTCTGTAACGGAATCCAAAGAGAGCAVTCTTGGTGAGTGC 1884
1499 euLysValLeuLeuHisSerMetAlaCysAsnGlnSerAlaValTyrLeu 1515
1885 TAAAGTGCTACTACACAGCATGGCCTGTAACCAAGTGACAGTTATCTA 1934
1516 GlnHisCysPheAlaThrGlnArgAlaLeuValSerLysPheProGluLe 1532
1935 CAACACTGTTTGGCTACACAGAGACCTTGGTTCAAGTTTCCTGAACT 1984
1532 uLeuPheGluGluGluTyrThrGluGlnCysAlaAspLeuCysLeuArgLeuL 1549
1985 CTTATTTGAAGAGAGACAGACAGAGTGCTGTGATTTATGCCCTCAGGCTTC 2034
1549 euArgHisCysSerSerSerIleGlyThrIleArgSerHisProSerAla 1565
2035 TCCGACACTGTAGCAGTAGCATCGGTACAATACGGTCAACCCACAGTGCC 2084
1566 SerLeuTyrLeuLeuMetArgGlnAsnPheGluIleGlyAsnAsnPheAl 1582
2085 TCCCTTTACCTACTAATGAGGCAAAACTTTGAGATTGGGAATTAACCTTGC 2134
1582 aaArgValLysMetGlnValPrometSerLeuSerSerLeuValGlyThrS 1599
2135 CAGGGTTAAAAATGCAGGTACCAATGTCACTATCCTCCTTGGTGGGCACAT 2184
1599 eArgLAsnPheAsnGluGluPheLeuArgArgSerLeuLysThrIleLeu 1615
2185 CTCAGAATTTTAATGAGAATTTCTTAAGACGTTCTCTTAAAGACTATATTG 2234
1616 ThrTyrAlaGluGluAspLeuGluLeuArgGluTyrThrPheProAspGl 1632
2235 ACATATGCTGAAGAGATCTGGAATTGAGGGAAACAACATTTCCTGATCA 2284
1632 nValGlnAspLeuValPheAsnLeuHisMetIleLeuSerAspThrValL 1649
2285 GGTCCAGGATCTGGTTTCAATCTCCATATGATTTCTTGATACTGTGA 2334
1649 ysmetLysGluHisGlnGluAspProGluMetLeuIleAspLeuMetTyr 1665
2335 AAATGAAGGAACACAGAGAGATCCTGAAATGTGATTGATCTAATGTAC 2384

1666 ArgIleAlaLysGlyTyrGlnThrSerProGlu...ArgLeuThrTrpLe 1681
2385 AGAATTGCCAAGGGTTTACCAGACCTCTCCAGATCTGGCATTTGACCTGTT 2434
1681 uGlnAsnMetAlaGlyLysHisSerGluArgSerAsnHisAlaGluAla 1698
2435 GCAGAACATGGCAGGCAAGCAGCTCAGAACGAAAGCAATCATCTGAAGCTG 2484
1698 laGlnCysLeuValHisSerAlaAlaLeuValAlaGluTyrLeuSerMet 1714
2485 CACAGTGTCTAGTCCACTCAGCAGCACTTGTGCTGAATATTGAGCATG 2534
1715 leuGluAspArgLysTyrLeuProValGlyCysValThrPheGlnAsnI 1731
2535 CTGAGAGACCGGAAATATCTTCCCTGGGATGTGTAACATTTCAAGATAT 2584
1731 eSerSerAsnValLeuGluGluSerAlaValSerAspValValSerP 1748
2585 TTCATCTAATGTTTTAGAAGAATCGCGGTCTCAGATGATGGTATCTC 2634
1748 roAspGluGluGlyIleCysSerGlyLysTyrPheThrGluSerGlyLeu 1764
2635 CAGATGAAGAAGGTATCTGCTCTGCAAAATACTTTACTGAGTCAGGACTT 2684
1765 ValGlyLeuLeuGluGlnAlaAlaAlaSerPheSerMetAlaGlyMetTy 1781
2685 GTGGGATTACTGGAACAAGCAGCTGCTTCTCTATAGCTGGCATGTA 2734
1781 rGluAlaValAsnGluValTyrLysValLeuIleProIleHisGluAla 1798
2735 TGAAGCAGTTAATGAAGTTTACAAGTACTTATTCCTATTCATGAAGCTA 2784
1798 snArgAspAlaLysLysLeuSerThrIleHisGlyLysLeuGlnGluAla 1814
2785 ATCGGATGCAAAAGAACTATCCCAATTCATGTGTAACCTTCAAGAAGCA 2834
1815 PheSerLysIleValHisGlnSerThrGlyTrpGluArgMetPheGlyTh 1831
2835 TTCAGCAAAATGTTTCATCAGAGTACTGGCTGGAGCGGATGTTGGCAC 2884
1831 rTyrPheArgValGlyPheTyrGlyThrLysPheGlyAspLeuAspGluG 1848
2885 CTATTTTCGTGTTGGTTTATGGAACCAAGTTCGGGGATTTGGATGAAC 2934
1848 lngluPheValTyrLysGluProAlaIleThrLysLeuAlaGluIleSer 1864
2935 AAGAAATTTGTTTACAAGAGAGCTGCAATTAACCAAACTTGCAGAGATATCT 2984
1865 HisArgLeuGluGlyPheTyrGlyGluArgPheGlyGluAspValValGl 1881
2985 CACAGATTGGAGGATTTTACGAGAAAGATTGGAGAGGATGTGTTGA 3034
1881 uValIleLysAspSerAsnProValAspLysCysLysLeuAspProAsnL 1898
3035 AGTAATCAAAAGACTCTAATCCTGTAGACAAAGTGAATTAGATCCCTAACA 3084
1898 ysaAlaTyrIleGlnIleThrTyrValGluProTyrPheAspThrTyrGlu 1914
3085 AGGCATATATTCAGATTACCTATGTGAGAGCCATACTTGACACATATGAG 3134
1915 metLysAspArgIleThrTyrPheAspLysAsnTyrAsnLeuArgArgPh 1931
3135 ATGAAGGACAGAATCACCTATTTCGACAAAAATTTACAATCTTTCGTGATT 3184
1931 eMetTyrCysThrProPheThrLeuAspGlyArgAlaHisGlyGluLeuH 1948
3185 CATGTACTGTACACCTTTACTTTAGATGGCCGTGCCATGGGGAACCTTC 3234
1948 isGluGlnPheLysArgLysThrIleLeuThrThrSerHisAlaPhePro 1964
3235 ATGAACAATTCAAAAGGAAGACATTCTGACTACGTCATGCCTTTCCT 3284

1965 TyrIleIeysThrArgValAsnValThrHisLysGluGluIleIleLeuTh 1981
|||||
3285 TATATATAAAACAAGGTCATATCTCATTAAGAAGAGATCATCTTAAC 3334
1981 rProIleGluValAlaIleGluAspMetGlnLysLysThrGlnGluLeuA 1998
|||||
3335 ACCAATTGAAGTTGCTATGTAGACATGCAGAAAAAGACACAGAGTTGG 3384
1998 lAPheAlaThrHisGlnAspProAlaAspProLysMetLeuGlnMetVal 2014
|||||
3385 CATTTGCACACATCAGATGCCGACAGACCCCAAAATGCTTCAGATGTA 3434
2015 LeuGlnGlySerValGlyThrThrValAsnGlnGlnGlyProLeuGluValAl 2031
|||||
3435 CTCCAGGATCTGTAGGACACACAGTGAATCAGGGGCTTTGGAAGTTGC 3484
2031 aGlnValPheLeuSerGluIleProSerAspProLysLeuPheArgHisH 2048
|||||
3485 CCAGGTTTTCTGCTGAATACCTAGTGACCCAAAGCTCTTCAGACATC 3534
2048 lAsnLysLeuArgLeuCysPheLysAspPheThrLysArgCysGluAsp 2064
|||||
3535 ATAAATAACTGCGACTCTGCTTAAAGATTTTACTAAAGGTGTGAAGAT 3584
2065 AlaleuArgLysAsnLysSerLeuIleGlyProValGlnLysGluTyrG1 2081
|||||
3585 GCCTTAGAAAAAATAAGAGCTTAATGGGCCGTTCAAAAGAGATATCA 3634

2081 nArgGluLeuGlyLysLeuSerSerPro 2090
|||||
3635 AAGGGAATTGGGGAACATATCTTCGCCCT 3662

seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA2001A.DAT:AAH41911

seq_documentation_block:

ID AAH41911 standard; cDNA; 4143 BP.

XX AAH41911;

XX 30-AUG-2001 (first entry)

XX Preliminary human CLASP-3 cDNA sequence Fig 1.

XX Human; cadherin-like asymmetry protein; CLASP, CLASP-3; immune response;
KW cell surface molecule; transmembrane protein; immunosuppressive; vaccine;
KW antiinflammatory; antiarthritic; antianaemic; dermatological; uropathic;
KW ophthalmological; antirheumatic; nephrotropic; antithyroid; antidiabetic;
KW neuroprotective; antiasthmatic; antibacterial; antisense therapy;
KW gene therapy; chromosome 1p31.1; ss.

XX Homo sapiens.

XX WO200142297-A2.

XX 14-JUN-2001.

XX 13-DEC-2000; 2000WO-US34171.

XX 13-DEC-1999; 99US-0170453.
PR 14-JAN-2000; 2000US-0176195.
PR 14-FEB-2000; 2000US-0182296.
PR 11-APR-2000; 2000US-0196267.
PR 11-APR-2000; 2000US-0196460.
PR 11-APR-2000; 2000US-0196527.
PR 11-APR-2000; 2000US-0196528.
PR 11-APR-2000; 2000US-0547276.
PR 13-OCT-2000; 2000US-0240503.
PR 13-OCT-2000; 2000US-0240508.

XX (ARBO-) ARBOR VITA CORP.

XX Lu P, Garman JD, Candia AF;

XX

DR WPI; 2001-375003/39.
DR P-PSDB; AAB99493.
XX
PT Novel isolated cadherin-like asymmetry protein (CLASP) -3, useful for
PT useful in preventing or treating a CLASP-3-mediated disease e.g.
PT autoimmune disease
XX
PS Example 5; Fig 1; 189pp; English.

XX The present invention describes an isolated polypeptide (I) comprising a
CC nucleotide (nt) sequence (S1a) that has at least 90% identity to the
CC sequence given in AAH41934 and is immunologically cross-reactive with
CC the derived amino acid (aa) sequence (S1b) given in AAB99495 or shares a
CC biological function with native CLASP-3. (I) has immunosuppressive,
CC antiinflammatory, antiarthritic, antianaemic, dermatological, uropathic,
CC ophthalmological, antirheumatic, nephrotropic, antithyroid, antidiabetic,
CC neuroprotective, antiasthmatic and antibacterial activities, and can be
CC used in antisense therapy, vaccine production and gene therapy. CLASP-3
CC related sequences can be used in preventing or treating a CLASP-3
CC mediated disease, preferably an autoimmune disease by inhibiting an
CC immune response. The autoimmune disease is caused or exacerbated by
CC increased activity of T helper 1 lymphocytes (TH1). AAH41893 to AAH41953
CC and AAB99491 to AAB99507 represent sequences which are used in the
CC exemplification of the present invention. CLASP-3 is localised in the
CC chromosome location 1p31.1.

XX Sequence 4143 BP; 1251 A; 829 C; 890 G; 1173 T; 0 other;

alignment_scores: Quality: 6157.50 Length: 1201
 Ratio: 5.131 Gaps: 1
Percent Similarity: 99.917 Percent Identity: 99.750

alignment_block:
US-09-737-246-2 x AAH41911 ..

Align seg 1/1 to: AAH41911 from: 1 to: 4143

891 LeuAsnArgSerArgSerLeuSerAsnSerAsnProAspIleSerGlyTh 907
:::|||||
4 GTAATCGTCTTCGAAGCCCTTAGTAATAGCAATCCAGATATATCTGGGAC 53
907 rProThrSerProAspAspGluValArgSerIleIleGlySerLysGlyL 924
|||||
54 TCCACAGTCACCAAGATGATGAAGTTTCGATCAATCATCGGAGTAAGGTT 103
924 euAspArgSerAsnSerTrpValAsnThrGlyGlyProLysAlaAlaPro 940
|||||
104 TAGATCGCTCCAAATCTCTGGGTTAACACTGTGTGCCAAAAGCTGCCCA 153
941 TrpGlySerAsnProSerProSerAlaGluSerThrGlnAlaMetAspAr 957
|||||
154 TGGGATCCAAACCCAGTCCCAAGTGCAGAAATCAACACAGGCTATGATCG 203
957 gSerCysAsnArgMetSerSerHisThrGluThrSerSerPheLeuGlnT 974
|||||
204 AAGTTGTAATCGTATGTCTTCGCACACAGACAGTCAGTTCTTACAAA 253
974 hrLeuThrGlyArgLeuProThrLysLysLeuPheHisGluGluLeuAla 990
|||||
254 CATTAACGGGACGCTTACCACTAAAAAGCTTTTTCACGAGAGAGCTGGCT 303
991 LeuGlnTrpValValCysSerGlySerValArgGluSerAlaLeuGlnG1 1007
|||||
304 TTGCAGTGGTGTGTTGCAGTGGCAGCGTTGGGAATCAGCTTTGCAACA 353
1007 nAlaTrpPhePheGluLeuMetValLysSerMetValHisHisLeuT 1024
|||||
354 AGCCTGGTCTTTTGAATTAATGTAAGAGCATGTGCACCAATTAT 403
1024 YrPheAsnAspLysLeuGluAlaProArgLysSerArgPheProGluArg 1040
|||||

404 ACITTAATGATAAACTTGAGGCTCCAAGAAAAGTCGTTTTCCAGAACGT 453
1041 PheMetAspSPIleAlaIaLeuValSerThrIleAlaSerAspIleVal 1057
|||||
454 TTCAATGGATTCAGATTCGACATCTGTGTGACGACGATTCGATGATATAGT 503
1057 lSerArgPheGlnLysAspThrGluMetValGluArgLeuAsnThrSerL 1074
|||||
504 TTCACGATTTTCAGAAGGACAGAAATGGTTGAGAGACTCAATACAAAGCC 553
1074 euAlaPhePheLeuAsnAspLeuLeuSerValMetAspArgGlyPheVal 1090
|||||
554 TTGCATTTCTTCTCAATGATCTGTGTCTGTATGACAGAGATTTGTT 603
1091 PheSerLeuIleLysSerCysTyrLysGlnValSerSerLysLeuTyrSe 1107
|||||
604 TTTAGCCTTATAAGTCCCTGCTATAAACAGGTGTCTTCAAAAGCTTTACTC 653
1107 rLeuProAsnProSerValLeuValSerLeuArgLeuAspPheLeuArgI 1124
|||||
654 ATTACCGAATCCAGTGTCTGTGTCCTTGAGGCTGGATTTTCTACGAA 703
1124 lIleLecysSerHisGluHisTyrValThrLeuAsnLeuProCysSerLeu 1140
|||||
704 TTCATCTGCATCATGAGCAGCATATGTACATTAAACCTTACCCTGCAGCTTA 753
1141 LeuThrProProAlaSerProSerProSerValSerSerAlaThrSerG 1157
|||||
754 CTACTCCACCTGCATCTCCATCACCTTCTGTCTCTGCAACATCTCA 803
1157 nSerSerGlyPheSerThrAsnValGlnAspGlnLysIleAlaAsnMetP 1174
|||||
804 GAGTTCGGATTTTCTACGAATGTACAAGACCAAAAGATTGCAATATGT 853
1174 heGluLeuSerValProPheArgGlnGlnHisTyrLeuAlaGlyLeuVal 1190
|||||
854 TTGAATTTATCCGTCCTTCCGCCAACAGCATATTGTCGACAGACTTGTG 903
1191 LeuThrGluLeuAlaValIleLeuAspProAspAlaGluGlyLeuPheG 1207
|||||
904 TTAACAGAGCTGGCTGTCAATTTTAGACCCTGATGCTGAAGGACTGTTTG 953
1207 yLeuHisLysLysValIleAsnMetValHisAsnLeuLeuSerSerHisA 1224
|||||
954 ATTGCATTAAGAAAGTCATCAATATGTACACAAATTACTCTCCAGTCACG 1003
1224 sPseAspProArgTyrSerAspProGlnIleLysAlaArgValAlaMet 1240
|||||
1004 ACTCAGACCCGCGGTACTCTGCACCCTCAGATAAAGGCTCGAGTGGCCATG 1053
1241 LeuTyrLeuProLeuIleGlyIleIleMetGluThrValProGlnLeuTy 1257
|||||
1054 TTGTATCTACCTCTGATTTGGTATTCATGGAACGTGTACTCAGCTGTA 1103
1257 rAspPheThrGluThrHisAsnGlnArgGlyArgProIleCysIleAlaT 1274
|||||
1104 TGATTTTACAGAAACTCACAAATCAACGAGGAAGACCAATTTGTATAGCCA 1153
1274 hrAspAspTyrGluSerGluSerGlySerMetIleSerGlnThrValAla 1290
|||||
1154 CTGATGATTAAGAAAGTGAGAGCGGAAGTATGATAAGCCAGACCGTTGCC 1203
1291 MetAlaIleAlaGlyThrSerValProGlnLeuThrArgProGlySerPh 1307
|||||
1204 ATGGCAATCGCAGGACATCGGTCCCTCAACTAACCAAGGCTGGCAGTTT 1253
1307 eLeuLeuThrSerThrSerGlyArgGlnHisThrThrPheSerAlaGluS 1324
|||||
1254 CCTCTCAGCTCAACGAGTGGCAGGCAACACTACTCTTTCAGCAGAAAT 1303
1324 erSerArgSerLeuLeuIleCysLeuLeuTrpValLeuLysAsnAlaAsp 1340
|||||
1304 CAAGTCGAAGCCTTTTGATCTGTCTACTTTGGGTTCTCAAAAATGCAGAT 1353

1341 GluThrValLeuGlnLysTrpPheThrAspLeuSerValLeuGlnLeuAs 1357
|||||
1354 GAAACAGTTCTACAGAAGTGGTTTACAGATCTCTCAGTCTTCAGCTAAA 1403
1357 narGLeuLeuAspLeuLeuTyrLeuCysValSerCysPheGluTyrLysG 1374
|||||
1404 CCGGCTATTAGATCTGCTTTATCTCTGTGTCTGTGCTTGTGAGTATAAAG 1453
1374 lYlLysLysValPheGluArgMetAsnSerLeuThrPheLysLysSerLys 1390
|||||
1454 GGAAAAAAGTGTGGAACGAATGAATAGCTTGACCTTTTAAGAAATCAAAA 1503
1391 AspMetArgAlaLysLeuGluGluAlaIleLeuGlySerIleGlyAlaAr 1407
|||||
1504 GACATGAGACCAAAAGCTTGAAGAAGCTATTCTTGGAGCAYAGGTGCCAG 1553
1407 gGlnGluMetValArgArgSerArgGlyGlnLeuGluArgSerProSerG 1424
|||||
1554 GCAAGAAATGGTACGGCGGAAGCCGAGACAGCTCGAGAGAACCCCATCTG 1603
1424 lYSerAlaPheGlySerGlnGluAsnLeuArgTrpArgLysAspMetThr 1440
|||||
1604 GAAGTGCCCTTTGGAAGTCAAGAAAATTTGAGGTGGAGGAAGATATGACT 1653
1441 HisTrpArgGlnAsnThrGluLysLeuAspLysSerArgAlaGluIleGl 1457
|||||
1654 CACTGGCGTCAAAACACAGAGAAGCTTGACAAATCAAGAGCAGAGATTGA 1703
1457 uHisGluAlaLeuIleAspGlyAsnLeuAlaThrGluAlaAsnLeuIleI 1474
|||||
1704 ACACGAAGCACTGATTTGATGGAACCTGGCTTACAGAAGCAAACTTAATCA 1753
1474 lLeuAspThrLeuGluIleValValGlnThrValSerValThrGluSer 1490
|||||
1754 TTTAGATACATTAAGATTTGTTGTTCAGACCCTTCTGTAAAGGAATCC 1803
1491 LysGluSerIleLeuGlyGlyValLeuLysValLeuLeuHisSerMetAl 1507
|||||
1804 AAAGAGAGCATCTTGGTGAGTGCCTAAAGTGCTACTACACAGCATGGC 1853
1507 acYsAsnGlnSerAlaValTyrLeuGlnHisCysPheAlaThrGlnArgA 1524
|||||
1854 CTGTAAACCAAGTGCAGTTTATCTACACACTGTTTGTCTACACAGAGAG 1903
1524 lAlaLeuValSerLysPheProGluLeuLeuPheGluGluThrGluGln 1540
|||||
1904 CCTGGTTTCAAAGTTTCTCTGAACCTTATTTGAAGAAAGACAGACAGCAG 1953
1541 CysAlaAspLeuCysLeuArgLeuLeuArgHisCysSerSerSerIleGl 1557
|||||
1954 TGTGCTGATTTATGCCCTCAGGCTTCTCGACACTGTAGCAGTAGCATCGG 2003
1557 yThrIleArgSerHisProSerAlaSerLeuTyrLeuLeuMetArgGlnA 1574
|||||
2004 TACAATACGGTCAACCCCAAGTGCCTCCCTTACTACTAATGAGGCAAA 2053
1574 snPheGluIleGlyAsnAsnPheAlaArgValLysMetGlnValPromet 1590
|||||
2054 ACTTTGAGATTTGGGAATAACTTTGCCAGGCTTAAAAATGCAGGTACCAATG 2103
1591 SerLeuSerSerLeuValGlyThrSerGlnAsnPheAsnGluGluPheLe 1607
|||||
2104 TCACTATCTCTCTGGTGGGACACATCTCGAATTTTAATGAAGAAATTCTT 2153
1607 uArgArgSerLeuLysThrIleLeuThrTyrAlaGluGluAspLeuGluL 1624
|||||
2154 AAGAGCTTCTTAAGACTATATATTGACATATGCTGAAGAAGATCTGGAAT 2203
1624 euArgGluThrThrPheProAspGlnValGlnAspLeuValPheAsnLeu 1640
|||||
2204 TGAGGGAACAACATTTCTCTGATCAGGTCCAGGATCTGTGTTTCAATCTC 2253

1641 HisMetIleuSerAspThrValLysMetLysGluHisGlnLysAspR 1657
|||||
2254 CATATGATCTTCTGTGACTGTGAATAAGAGAACACAGAGATCC 2303
1657 oGluMetLeuIleAspLeuMetTyrArgIleAlaLysGlyTyrGlnThrs 1674
|||||
2304 TGAATGTGATTGATCTAATGTACAGAAATGGCCAAGGTTACAGACCT 2353
1674 erProGlu...ArgLeuThrTyrLeuGlnAsnMetAlaGlyLysHisSer 1689
|||||
2354 CTCAGATCTGGGATGTGACCTGGTGGACAGACATGGCAGGCAACACTCA 2403
1690 GluArgSerAsnHisAlaGluAlaAlaGlnCysLeuValHisSeraAlaI 1706
|||||
2404 GAACGAAGCATCATGCTGAAGCTGCACAGTGTCTAGTCCACTCAGCAGC 2453
1706 aLeuValAlaGluTyrLeuSerMetLeuGluAspArgLysTyrLeuPro 1723
|||||
2454 ACTGTGCTGAATATTGTAGCATGCTGGAGACCGGAATATCTTCTCTG 2503
1723 aIGlyCysValThrPheGlnAsnIleSerSerAsnValLeuGlnGluSer 1739
|||||
2504 TGGATGTGTAACTTTGAGATATTCATCTAATGTTTGAAGAATCT 2553
1740 AlavalSerAspAspValValSerProAspGluGluGlyIleCysSerG 1756
|||||
2554 GCGGTCTCAGATGATGTGATCTCAGATGAAGAAGGTATCTGCTCTGG 2603
1756 yLysTyrPheThrGluSerGlyLeuValGlyLeuLeuGlnAlaAlaA 1773
|||||
2604 AAAATACTTACTGAGTCAGACCTGTGGGATTACTGGAACAAGCAGCTG 2653
1773 laSerPheSerMetAlaGlyMetTyrGluAlaValAsnGluValTyrLys 1789
|||||
2654 CTTCCTTCTCTATGGCTGGCATGTATGAAGCAGTTAATGAAGTTTACAA 2703
1790 ValLeuIleProIleHisGluAlaAsnArgAspAlaLysLysLeuSerTh 1806
|||||
2704 GTACTTATTCCTATTTCATGAAGCTAATCGGGATGCAAAAGAACTATCCAC 2753
1806 rIleHisGlyLysLeuGlnGluAlaPheSerLysIleValHisGlnSerT 1823
|||||
2754 AATTCAATGTAACCTTCAGAAGCAATTCAGCAAAATGTTCATCAGAGTA 2803
1823 hrGlyTyrPGLuArgMetPheGlyThrTyrPheArgValGlyPheTyrGly 1839
|||||
2804 CTGGCTGGGAGCGGATGTTGGCACTATTTCGTGTTGTTTATGGA 2853
1840 ThrLysPheGlyAspLeuAspGluGlnGluPheValTyrLysGluProAl 1856
|||||
2854 ACCAAGTTCGGGATTTGGATGAGACAAGAATTGTTACAGAAGAGCCTGC 2903
1856 aIleThrLysLeuAlaGluIleSerHisArgLeuGluGlyPheTyrGlyG 1873
|||||
2904 AATAACCAAACTTGCAGAGATATCTCACAGATTGGAGGGAATTTTACGGAG 2953
1873 LuArgPheGlyGluAspValValGluValIleLysAspSerAsnProVal 1889
|||||
2954 AAAGATTGGAGAGGATGTGGTGAAGTAATCAAGACTCTAATCCTGTA 3003
1890 AspLysCysLysLeuAspProAsnLysAlaTyrIleGlnIleThrTyrVa 1906
|||||
3004 GACAAGTGTAAATTAGATCCTAACAGGCATATATTCAGATTACCTATGT 3053
1906 IGLuProTyrPheAspThrTyrGluMetLysAspArgIleThrTyrPhea 1923
|||||
3054 GGAGCCATACTTTGACACATATGAGATGAAGACAGAAATCACTTATTCG 3103
1923 sPLysAsnTyrAsnLeuArgArgPheMetTyrCysThrProPheThrLeu 1939
|||||
3104 ACAAAAATTACAATCTTCGTGATTCATGTACTGTACACCCCTTACTTTA 3153
1940 AspGlyArgAlaHisGlyLuleuHisGluGlnPheLysArgLysThrI 1956

|||||
3154 GATGGCCGTGCCCATGGGGAACCTTCATGAACAATTCAAAAGAGACCAT 3203
1956 eLeuThrThrSerHisAlaPheProTyrIleLysThrArgValAsnValT 1973
|||||
3204 TCTGACTACGTCATATGCCCTTTCCTTATATTAAACAAGGTTCAATGTCA 3253
1973 hrHisLysGluGluIleIleLeuThrProIleGluValAlaIleGluasp 1989
|||||
3254 CTCATAAAGAAGAGATCATCTTAACACCAATTGAAGTTGCTATGAGAGAC 3303
1990 MetGlnLysLysThrGlnGlnLuleuAlaPheAlaThrHisGlnAspProAl 2006
|||||
3304 ATGCAGAAAAAGACACAGAGATTGGCATTTGCCAACACATCAGGATCCCGC 3353
2006 aAspProLysMetLeuGlnMetValLeuGlnGlySerValGlyThrThrv 2023
|||||
3354 AGACCCCAAAATGCTTCAGATGGTACTCCAGGGATCTGTAGGCACACAG 3403
2023 aLasnGlnGlyProLeuGluValAlaAlaGlnValPheLeuSerGluIlePro 2039
|||||
3404 TGAATCAGGGGCTTTTGGAAAGTTGCCAGGTTTCTGTCTGAATAACTCT 3453
2040 SerAspProLysLeuPheArgHisHisAsnLysLeuArgLeuCysPheLy 2056
|||||
3454 AGTGACCCAAAGCTCTTCAGACATCATATTAACCTGCGACTCTGCTTTAA 3503
2056 sasPheThrLysArgCysGluAspAlaLeuArgLysAsnLysSerLeuI 2073
|||||
3504 AGATTTTACTAAAGGCTGTGAAGATGCCCTTAAGAAAAAATAAGACTTAA 3553
2073 legIyProValGlnLysGluTyrGlnArgGluLeuGlyLysLeuSerSer 2089
|||||
3554 TTGGCGCGGTTTCAAAAGAGATATCAAAAGGAATTGGGAAACTATCTTCG 3603
2090 Pro 2090
|||
3604 CCT 3606
seq_name: /SIDSl/gcgdata/hold-geneseq/geneseqn-emb1/NA2001A.DAT:AAH41912
seq_documentation_block:
ID AAH41912 standard; cDNA; 4144 BP.
XX AAH41912;
AC
XX
DT 30-AUG-2001 (first entry)
DE
XX Human CLASP-3 cDNA sequence Fig 4.
XX
KW Human; cadherin-like asymmetry protein; CLASP; CLASP-3; immune response;
KW cell surface molecule; transmembrane protein; immunosuppressive; vaccine;
KW antiinflammatory; antiarthritic; antianaemic; dermatological; uropathic;
KW ophthalmological; antirheumatic; nephrotropic; antidiabetic;
KW neuroprotective; antiasthmatic; antibacterial; antisense therapy;
KW gene therapy; chromosome 1p31.1; ss.
XX Homo sapiens.
OS
XX
PN WO200142297-A2.
XX
PD 14-JUN-2001.
XX
PF 13-DEC-2000; 2000WO-US34171.
XX
PR 13-DEC-1999; 99US-0170453.
PR 14-JAN-2000; 2000US-0176195.
PR 14-FEB-2000; 2000US-0182296.
PR 11-APR-2000; 2000US-0196267.
PR 11-APR-2000; 2000US-0196460.
PR 11-APR-2000; 2000US-0196527.
PR 11-APR-2000; 2000US-0196528.
PR 11-APR-2000; 2000US-0547276.

PR 13-OCT-2000; 2000US-0240503.
PR 13-OCT-2000; 2000US-0240508.
XX
PA (ARBO-) ARBOR VITA CORP.
XX
PI Lu P, Garman JD, Candia AF;
XX
DR WPI; 2001-375003/39.
DR P-PSDB; AAB99494.
XX
PT Novel isolated cadherin-like asymmetry protein (CLASP) -3, useful for
PT useful in preventing or treating a CLASP-3-mediated disease e.g.
PT autoimmune disease
XX
PS Disclosure; Fig 4A; 189pp; English.
XX
CC The present invention describes an isolated polypeptide (I) comprising a
CC nucleotide (nt) sequence (Sia) that has at least 90% identity to the
CC sequence given in AAH41934 and is immunologically cross-reactive with
CC the derived amino acid (aa) sequence (S1b) given in AAB99495 or shares a
CC biological function with native CLASP-3. (I) has immunosuppressive,
CC antiinflammatory, antirheumatic, antianaemic, dermatological, uropathic,
CC ophthalmological, antirheumatic, nephrotropic, antithyroid, antidiabetic,
CC neuroprotective, antiasthmatic and antibacterial activities, and can be
CC used in antisense therapy, vaccine production and gene therapy. CLASP-3
CC related sequences can be used in preventing or treating a CLASP-3
CC mediated disease, preferably an autoimmune disease by inhibiting an
CC immune response. The autoimmune disease is caused or exacerbated by
CC increased activity of T helper 1 lymphocytes (TH1). AAH41893 to AAH41953
CC and AAB99491 to AAB99507 represent sequences which are used in the
CC exemplification of the present invention. CLASP-3 is localised in the
CC chromosome location 1p31.1.
XX
SQ Sequence 4144 BP; 1252 A; 829 C; 890 G; 1173 T; 0 other;

alignment_scores:
Quality: 6157.50 Length: 1201
Ratio: 5.131 Gaps: 1
Percent Similarity: 99.917 Percent Identity: 99.750

alignment_block:

US-09-737-246-2 x AAH41912 ..

Align seg 1/1 to: AAH41912 from: 1 to: 4144

891 LeuAsnArgSerArgSerLeuSerAsnSerAsnProAspIleSerGlyTh 907
:::|||||
5 GTAATCGTCTCGAAGCCTTAGTAATAGCAATCCAGATATATCTGGGAC 54
907 rProThrSerProAspAspGluValArgSerIleIleGlySerGlyL 924
|||||
55 TCCACGTCACAGATGATGAAGTTCGATCAATCATCGGAGTAAGGTT 104
924 euAspArgSerAsnSerTrpValAsnThrGlyGlyProLysAlaAlaPro 940
|||||
105 TAGATCGCTCCAATTCTCTGGTTAACTGGTGTCCAAAAGCTGCCCA 154
941 TrpGlySerAsnProSerProSerAlaGluSerThrGlnAlaMetAsp 957
|||||
155 TGGGATCCAAACCCAGTCCAAAGTCAGAAATCAACACAGCGTATGATCG 204
957 gSerCysAsnArgMetSerSerHisThrGluThrSerSerPheLeuGln 974
|||||
205 AAGTGTATTCGTATGTCTTCGCACACAGAGAGCGTCAAGTTCTTACAAA 254
974 hrLeuThrGlyArgLeuProThrLysLysLeuPheHisGluGluLeuAla 990
|||||
255 CATTAACGGGAGCGCTTACCAACTAAAAAGCTTTTTCACGAGGAGCGTGGCT 304
991 LeuGlnTrpValValCysSerGlySerValArgGluSerAlaLeuGlnG 1007
|||||
305 TTGCAGTGGGTTGTTTGCAGTGGCAGCGTTTCGGGAATCAGCTTTGCAACA 354

1007 nAlaTrpPhePheGluLeuMetValLysSerMetValHisHisLeu 1024
|||||
355 AGCCTGGTCTTTTGAATTAATGTTAAAGACATGTTGCACCATTTAT 404
1024 yrPheAsnAspLysLeuGluAlaProArgLysSerArgPheProGluArg 1040
|||||
405 ACTTTAATGATAACTTGAGGCTCCAAAGGAAAGTCGTTTCCAGAACGT 454
1041 pheMetAspAspIleAlaAlaLeuValSerThrIleAlaSerAspIleVal 1057
|||||
455 TTCATGGATGACATTTGCAGCTCTTGTTCAGCACGATTGCTAGTATATAGT 504
1057 lSerArgPheGlnLysAspThrGluMetValGluArgLeuAsnThrSerL 1074
|||||
505 TTCACGATTTCAGAGAGACACAGAAATGGTTGAGAGACTCAATACAAAGCC 554
1074 euAlaPhePheLeuAsnAspLeuLeuSerValMetAspArgGlyPheVal 1090
|||||
555 TTGCATTTCTTCTCAATGATCTGTGTCTGTATGACAGAGAGATTTGTT 604
1091 pheSerLeuIleLysSerCysTyrLysGlnValSerSerLysLeuTyrSe 1107
|||||
605 TTTAGCCTTATTAAGTCTCTGCTATAAACAGGTGCTTCAAGCTTTACTC 654
1107 rLeuProAsnProSerValLeuValSerLeuArgLeuAspPheLeuArgI 1124
|||||
655 ATTACCGAATCCAGTGTCTGTGTGTCTTGAGGCGCTGGATTTTCTACGAA 704
1124 leIleCysSerHisGluHisTyrValThrLeuAsnLeuProCysSerLeu 1140
705 TCATCTGCAGTCATGAGCACATATGTTACATTAACTTACCTGCAGCTTA 754
1141 LeuThrProProAlaSerProSerProSerValSerSerAlaThrSerG 1157
|||||
755 CTTACTCCACCTGCATCTCCATCACCTTCTGTTCTTCTGCAACATCTCA 804
1157 nSerSerGlyPheSerThrAsnValGlnAspGlnLysIleAlaAsnMetP 1174
|||||
805 GAGTCTCTGATTTTCTACGAATGTACACAGACCAAAAGATTGCCAATATGT 854
1174 heGluLeuSerValProPheArgGlnGlnHisTyrLeuAlaGlyLeuVal 1190
|||||
855 TTGAATTATCCGTGCTTTCGCCCAACAGCATTTATTTGGCAGGACTTGTG 904
1191 LeuThrGluLeuAlaValIleLeuAspProAspAlaGluGlyLeuPheG 1207
|||||
905 TTAACAGAGCTGGCTGCTCATTTTAGACCCGTGATGCTGAAGGACTGTTGG 954
1207 yLeuHisLysLysValIleAsnMetValHisAsnLeuLeuSerSerHisA 1224
|||||
955 ATTGCATAAGAAAGTCATCAATATGTGTACACAAATTACTCTCCAGTACG 1004
1224 spSerAspProArgTyrSerAspProGlnIleLysAlaArgValAlaMet 1240
|||||
1005 ACTCAGACCCCGGCTACTGTGACCCCTCAGATTAAGGCTCGAGTGGCCATG 1054
1241 LeuTyrLeuProLeuIleGlyIleIleMetGluThrValProGlnLeuTy 1257
|||||
1055 TTGTATCTACCTCTGTGATTGGTATTATCATGGAACCTGTACCTCAGCTGTA 1104
1257 rAspPheThrGluThrHisAsnGlnArgGlyArgProIleCysIleAlaT 1274
|||||
1105 TGATTTTACAGAACTCACAATCAACGAGGAGAGACCAATTGTATAGCCA 1154
1274 hrAspAspTyrGluSerGluSerGlySerMetIleSerGlnThrValAla 1290
|||||
1155 CTGATGATTATGAAGTGAGAGCGGAAGTATGATTAAGCCAGACCGTTGCC 1204
1291 MetAlaIleAlaGlyThrSerValProGlnLeuThrArgProGlySerPh 1307
|||||
1205 ATGCAATCGCAGGAGACATCGGTCCCTCAACTAACAAAGGCGCTGGCAGCTT 1254

1307 eLeuLeuThrSerThrSergLysArgGlnHisThrThrPheSerAlaGlns 1324
|||||
1255 CCTCCTCACGTCACAGAGTGGCAGGCAACACACTACCTTTTCAGCAGAAAT 1304
1324 erSerArgSerLeuLeuIleCysLeuLeuTrpValLeuLysAsnAlaasp 1340
|||||
1305 CAAGTCGAAGCCTTTTGATCTGCTCTACTTTGGGTTCTCAAAAAATGCAGAT 1354
1341 GluThrValLeuGlnLysTrpPheThrAspLeuSerValLeuGlnLeuas 1357
|||||
1355 GAAACAGTTTACAGAGAGTGGTTTACAGATCTCTCAGTCTTGACAGCTAAA 1404
1357 nArgLeuLeuAspLeuLeuTyrLeuCysValSerCysPheGluTyrLysG 1374
|||||
1405 CCGGCTATTAGATCTGCTTATCTCTGTGTCTTGCTTGAGTATATAAG 1454
1374 LylLysLysValPheGluArgMetAsnSerLeuThrPheLysLysSerLys 1390
|||||
1455 GGAATAAAAGTGTTGTAACGAATGAATAGCTTGACCTTTAAGAAATCAAAA 1504
1391 AspMetArgAlaLysLeuGluGluAlaIleLeuGlySerIleGlyAlaAr 1407
|||||
1505 GACATGAGAGCAAAAGCTTGAAGAAGCTATTCTTGAGCATAGTGCCAG 1554
1407 gGlnGluMetValArgArgSerArgGlyGlnLeuGluArgSerProserg 1424
|||||
1555 GCAAGAAATGTTACGGCGGAAGCCGAGACAGCTCGAGAGAAGCCCATCTG 1604
1424 LysSerAlaPheGlySergGlnGluAsnLeuArgTrpArgLysAspMetThr 1440
|||||
1605 GAAGTGCCTTGGAAAGTCAAGAAATTTGAGGTGAGAGAAGATATGACT 1654
1441 HisTrpArgGlnAsnThrGluLysLeuAspLysSerArgAlaGluIleGl 1457
|||||
1655 CACTGGCGTCAAAACACAGAGAAGCTTGACAAATCAAGAGCAGAGATTGA 1704
1457 uHisGluAlaLeuIleAspGlyAsnLeuAlaThrGluAlaAsnLeuIleI 1474
|||||
1705 ACACGAAGCAGCTGATTTGATGAAACCTGGCTACAGAGAACAACCTAATCA 1754
1474 leLeuAspThrLeuGluIleValValGlnThrValSerValThrGluSer 1490
|||||
1755 TTTTAGATACATTAGAGATTGTTGTTTCAGACCCTTCTGTAAACGGAATCC 1804
1491 LysGluSerIleLeuGlyGlyValLeuLysValLeuLeuHisSerMetAl 1507
|||||
1805 AAAGAGAGCATCTTGGTGAGTGTCTAAAGTGTCTACTACACAGCATGGC 1854
1507 acYsAsnGlnSerAlaValTyrLeuGlnHisCysPheAlaThrGlnArgA 1524
|||||
1855 CTGTAACCAAGTGCAGTTTATCTACACACACTGTTTGTCTACACAGAGAG 1904
1524 lAleuValSerLysPheProGluLeuLeuPheGluGluTyrThrGluGln 1540
|||||
1905 CCTTGGTTTCAAGTTTCCCTGACACTTTATTGGAAGAGAAGACAGACAG 1954
1541 CysAlaAspLeuCysLeuArgLeuLeuArgHisCysSerSerSerIleGl 1557
|||||
1955 TGTGCTGATTATTCCTCAGGCTTCTCCGACACTGTAGCAGTAGCATCGG 2004
1557 yThrIleArgSerHisProSerAlaSerLeuTyrLeuLeuMetArgGlnA 1574
|||||
2005 TACAATACGGTCACACCCTCAGTGCCTCCCTTACCTACTAATGAGGCCAAA 2054
1574 snPheGluIleGlyAsnAsnPheAlaArgValLysMetGlnValPromet 1590
|||||
2055 ACTTTGAGATTGGGAATTACTTTCAGGGGTTAAATGCAGGTATCCAATG 2104
1591 SerLeuSerSerLeuValGlyThrSergGlnAsnPheAsnGluGluPheLe 1607
|||||
2105 TCACATATCCTCCTTGGTGGGCACATCTCAGAATTTTAATGAGAATTTCTT 2154
1607 uArgArgSerLeuLysThrIleLeuThrTyrAlaGluGluAspLeuGluL 1624

|||||
2155 AAGACGTTCTCTAAAGACTATATTTGACATATCTGAAGAAGATCTGGAAT 2204
1624 euArgGluThrThrPheProAspGlnValGlnAspLeuValPheAsnLeu 1640
|||||
2205 TGAGGGAAACAACATTTCTGATCGAGTCCAGGATCTGGTTTCAATCTC 2254
1641 HisMetIleLeuSerAspThrValLysMetLysGluHisGlnGluAspPr 1657
|||||
2255 CATATGATTTCTTCTGATCTACTGTGAAAAATGAAGGAACACAGAGGATCC 2304
1657 oGluMetLeuIleAspLeuMetTyrArgIleAlaLysGlyTyrGlnThrS 1674
|||||
2305 TGAATGCTTGATTGATCTAATGTACAGAATTGCCAAGGGTTTACCAGACCT 2354
1674 erProGlu...ArgLeuThrTrpLeuGlnAsnMetAlaGlyLysHisSer 1689
|||||
2355 CTCACGATCTGCGGATTCGACCTGGTTGCAGAACATGGCAGCAAGCAGCTCA 2404
1690 GluArgSerAsnHisAlaGluAlaAlaGlnCysLeuValHisSerAlaAl 1706
|||||
2405 GAACGAAGCAATCATGCTGAAGCTGCACAGATGTCCTAGTCCACTCAGCAGC 2454
1706 AleuValAlaGluTyrLeuSerMetLeuGluAspArgLysTyrLeuProV 1723
|||||
2455 ACTTGTCTGTAATATTGAGCATGCTGAGAGACCGGAATATCTTCCCTG 2504
1723 aGlyCysValThrPheGlnAsnIleSerSerAsnValLeuGluGluSer 1739
|||||
2505 TGGGATGTGTAACATTTCAAGATATTTCAATCAATGTTTAGAAGAAATCT 2554
1740 AlavalSerAspAspValValSerProAspGluGluGlyIleCysSergL 1756
|||||
2555 GCGGCTCTCAGATGATGTGTATCTCCAGATGAGAAGATATCTGCTCGG 2604
1756 yLysTyrPheThrGluSerGlyLeuValGlyLeuLeuGluGlnAlaAla 1773
|||||
2605 AAAATACTTTACTGAGTCAGGACTGTGTGGGATTACTGGAACAAGCAGCTG 2654
1773 lAserPheSerMetAlaGlyMetTyrGluAlaValAsnGluValTyrLys 1789
|||||
2655 CTTCCTTCTCTATGCGCTGGCATGTATGAAGCAGTTAATGAAGTTTACAAA 2704
1790 ValLeuIleProIleHisGluAlaAsnArgAspAlaLysLysLeuSerTh 1806
|||||
2705 GTACTTATTCCTATTCATGAAGCTAATPCGGATGCAAAAGAACTATCTCAC 2754
1806 rIleHisGlyLysLeuGlnGluAlaPheSerLysIleValHisGlnSerT 1823
|||||
2755 AATTCAATGGTAACCTTCAAGAAGCATTCAGCAAAATGTTTCATCAGAGTA 2804
1823 hrcLysTrpGluArgMetPheGlyThrTyrPheArgValGlyPheTyrGly 1839
|||||
2805 CTGGCTGGGAGCGGATGTTTGGCACCTATTTTCGTGTTTATGGA 2854
1840 ThrLysPheGlyAspLeuAspGluGlnGluPheValTyrLysGluProAl 1856
|||||
2855 ACCAAGTTCGGGGGATTTGGATGAACAAGAAATTTGTTTACAAGGAGCCCTGC 2904
1856 aileThrLysLeuAlaGluIleSerHisArgLeuGluGlyPheTyrGlyG 1873
|||||
2905 AATAACCAAACTTGCAGAGATATCTCACAGATTGGAGGATTTTACGGAG 2954
1873 LuArgPheGlyGluAspValValGluValIleLysAspSerAsnProVal 1889
|||||
2955 AAAGATTTGGAGAGGATGTGTTGAAGTAAATCAAAAGACTTAATCCTGTA 3004
1890 AspLysCysLysLeuAspProAsnLysAlaTyrIleGlnIleThrTyrVa 1906
|||||
3005 GACAAGTGTAAATATAGATCTTAACAAGCGCATATATTCAGATTTACTATGT 3054
1906 lGluProTyrPheAspThrTyrGluMetLysAspArgIleThrTyrPheA 1923
|||||

3055 GGAGCCATACTTTGACACATATGAGATGAAGACAGAAATCACCCTATTTCG 3104
1923 sPlYsAsnTyTrAsnLeuArGArGpHeMeTyTCysThrProPheThrLeu 1939
|||||
3105 ACAAAATTTACAATCTTCGTCGATTCATGTACTGTACACCCCTTACTTTTA 3154
1940 AspGLYArGAlAHISGLYGLULeuHISGLInPhelYsArGLYsThrIL 1956
|||||
3155 GATGGCCGTGCCCATGGGGAACTTCATGAACAATTCAAAAGGAAGACCAT 3204
1956 eleUThrThrSerHISAlaPheProTyrlleLYsThrArgValAsnValT 1973
|||||
3205 TCTGACTACGTTCATGCCTTTCCTATATTAAACAAGGTCATGTCA 3254
1973 hrHISLYsGLULeUleUleUleUThrProILeGLUValAlaILeGLUAsp 1989
|||||
3255 CTCATAAGAAGAGATCATCTTAACACCAATTGAAGTTGCTATTGAGGAC 3304
1990 MetGLInLYsThrGLInGLULeuAlaPheAlaThrHISGLInAspProAl 2006
|||||
3305 ATGCAGAAAAAGACACAGGAGTTGGCATTTGCCAACACATCAGGATCCGC 3354
2006 aAsPProLYsMeLeUGInMeTValLeUGInGLYSerValGLYThrThrv 2023
|||||
3355 AGACCCCAAAATGCTTCAGATGGTACTCCAGGATCTGTAGCACACACAG 3404
2023 aLAsnGLInGLYProLeUGLUValAlaGLInValPheLeuSerGLULePro 2039
|||||
3405 TGAATCAGGGCCCTTGGAAAGTTGCCCGCTTTTCTGTCTGAATAACCT 3454
2040 SerAsPProLYsLeUPheArGHISHisAsnLYsLeuArgLeUCysPhely 2056
|||||
3455 AGTGACCCCAAGCTCTTCAGACATCATATAAACTGCGACTCTGCTTTAA 3504
2056 sAsPpHeThrLYsArGCysGLUAspAlaLeuArGLYsAsnLYsSerLeuI 2073
|||||
3505 AGATTTACTTAAAGGTGTGAGATGCCCTTAAGAAAAATAAGAGCTTAA 3554
2073 leGLYProValGLInLYsGLUTyrGLInArGLULeUGLYLYsLeuSerSer 2089
|||||
3555 TTGGGCCCGTTCAAAAGAGATATCAAGGGAATTGGGGAACCTATCTTCG 3604
2090 Pro 2090
|||
3605 CCT 3607
seq_name: /SIDSl/gcgydata/hold-geneseq/geneseqn-emb1/NA2001A.DAT:AAS07373
seq_documentation_block:
ID AAS07373 standard; cDNA; 7215 BP.
XX AAS07373;
AC AAS07373;
XX 26-SEP-2001 (first entry)
DT 26-SEP-2001 (first entry)
XX Human cDNA encoding CLASP-5.
DE Human cDNA encoding CLASP-5.
XX Human; CLASP-5; cadherin-like asymmetry protein; immune gateway;
KW immunogen; antibody; autoimmune disease; rheumatoid arthritis;
KW multiple sclerosis; leukaemia; insulin dependent diabetes mellitis;
KM acquired immunodeficiency syndrome; AIDS; ss.
XX
OS Homo sapiens.
XX
FH Key location/Qualifiers
FT CDS 112..6048
FT FT /*tag= a
FT FT /product= "CLASP-5"
FT FT 1498..1519
FT FT primer_bind /*tag= b
FT FT /note= "PCR primer HC5gs1"
FT FT primer_bind 3087..3113
FT FT /*tag= c

FT primer_bind /note= "Antisense oligonucleotide AAS07384"
FT 3087..3113
FT /*tag= d
FT primer_bind /note= "Antisense oligonucleotide AAS07384"
FT 3232..3259
FT /*tag= e
FT primer_bind /note= "Antisense oligonucleotide AAS07384"
FT complement (3642..3660)
FT /*tag= f
FT primer_bind /note= "PCR primer HC5AS10b"
FT 4728..4756
FT /*tag= g
FT /note= "Antisense oligonucleotide AAS07384"
XX
PN WO200142296-A2.
XX
XX 14-JUN-2001.
PD
XX 13-DEC-2000; 2000WO-US34163.
PF
XX 13-DEC-1999; 99US-0170453.
PR 14-JAN-2000; 2000US-0176195.
PR 14-FEB-2000; 2000US-0182296.
PR 11-APR-2000; 2000US-0196267.
PR 11-APR-2000; 2000US-0196460.
PR 11-APR-2000; 2000US-0196527.
PR 11-APR-2000; 2000US-0196528.
PR 11-APR-2000; 2000US-0547276.
PR 13-OCT-2000; 2000US-0240503.
PR 13-OCT-2000; 2000US-0240508.
PR 13-OCT-2000; 2000US-0240543.
PR 13-OCT-2000; 2000US-0240539.
XX
PA (ARBOR-) ARBOR VITA CORP.
XX
PI Lu P, Garman JD, Candia AF;
XX
DR WPI; 2001-367865/38.
DR P-PSDB; AAU04024.
XX
XX CLASP-5 polynucleotides, proteins and antibodies are used to prevent or
PT treat a CLASP-5 mediated disease, such as an autoimmune disease e.g.
PT Rheumatoid arthritis -
XX
PS Claim 1; Fig 6; 188bp; English.
XX
CC The sequence encodes a cadherin-like asymmetry protein, CLASP-5, which
CC is a transmembrane protein of the immune system involved in the formation
CC of the immune gateway. CLASP-5, polynucleotides encoding it and an
CC anti-CLASP-5 antibody are used to prevent or treat a CLASP-5 mediated
CC disease, such as an autoimmune disease caused or exacerbated by
CC increased activity of TH1 cells. These diseases may include Addison's
CC disease, rheumatoid arthritis, dermatitis, Multiple sclerosis, systemic
CC Lupus Erythematosus and Autoimmune Thyroiditis, inflammatory conditions
CC (e.g. ischaemia-reperfusion) and responses, Leukaemia, acquired
CC immunodeficiency syndrome, AIDS, syphilis, typhoid, Burkitt's lymphoma,
CC Grave's disease and insulin dependent diabetes mellitis. CLASP-5 is used
CC for inhibiting an immune response in a cell (T cell or B cell) by
CC interfering with the expression of a CLASP-5 gene in the cell, the
CC ability of a CLASP-5 protein to bind to another cell or the ability of a
CC CLASP-5 protein to bind to another protein. The antibody binding CLASP-5
CC is used to inhibit an immune response is a subject. The polynucleotide is
CC used to detect CLASP-5 expression in cells and for diagnosis of
CC diseases and disorders associated with aberrant expression of CLASP-5.
XX
SQ Sequence 7215 BP; 2005 A; 1734 C; 1721 G; 1755 T; 0 other;

alignment_scores: length: 2071
 quality: 5785.50 Gaps: 25
 Ratio: 3.487
Percent Similarity: 80.106 Percent Identity: 55.722

alignment_block:

US-09-737-246-2 x AAS07373 ..

Align seg 1/1 to: AAS07373 from: 1 to: 7215

```
44 SerHisHisThrThrValProLeuThr.....GluAlaVala 55
   :::::::::::||||| ||::: ||
34 ACGGCCGCTGAAGTGAATCTTCAAGCCTCAGTTTATGACCCCTGT 83
   :::::::::::||||| ||::: ||
55 IAspProValAspLeuGluAspTyrLeuIleThrHisProLeuAlaVala 72
   :::::::::::||||| ||::: ||
84 GGAGCCAGTGGACTTGAAGAGACTTCTGATGACACACCTGAACAGCCTGG 133
   :::::::::::||||| ||::: ||
72 sPserGlyProLeuArgAspLeuIleGluPheProProAspPileGlu 88
   :::::::::::||||| ||::: ||
134 ATGTGCAGCTTGGCCAGAGCTCGGGAGCTTCACCTGATGACGACTTGGAC 183
   :::::::::::||||| ||::: ||
89 ValValTyrSerProArgAspCysArgThrLeuValSerAlaValProGlu 105
   :::::::::::||||| ||::: ||
184 GTGGTGTTCACGCCCAAGAAATGATGAGACTTGCAGCCCTCTTGGCCGGA 233
   :::::::::::||||| ||::: ||
105 uGluSer..GluMetAspProHisValArgAspCysIleArgSerTyrT 121
   :::::::::::||||| ||::: ||
234 GGAAGGGGTGAACCTGACCCCTCATGTACAGGACTGTGTTCAGACCTACA 283
   :::::::::::||||| ||::: ||
121 hrGluAspTrpAlaIleValIleArgLys..... 130
   :::::::::::||||| ||::: ||
284 TCCGTGAGTGGCTAATGCTGAACCGAAAAACCAAGGAAGTCCAGAAATC 333
   :::::::::::||||| ||::: ||
131 .....TyrHisLysLeuGlyThrGlyPheAsnProAsnThrLeuAspLys 145
   :::::::::::||||| ||::: ||
334 TGTGGCTTTAAAAAGACTGATCT..... 357
   :::::::::::||||| ||::: ||
145 sGlnLysGluArgGlnLysGlyLeuProLysGlnValPheGluSerAspG 162
   :::::::::::||||| ||::: ||
358 .CGAAAAGATTTTACACAGACGCTTCCGAAACAGACGTTTGAGTCGGA 406
   :::::::::::||||| ||::: ||
162 LuAlaProAspGlyAsnSerTyrGlnAspAspGlnAspAspLeuLysArg 178
   :::::::::::||||| ||::: ||
407 CCTTGGAG.....TGCAGTGAACCCGCTGCTCAGGCAAGCCCCGCCAC 450
   :::::::::::||||| ||::: ||
179 ArgSerMetSerIleAspAspThrProArgGlySerTrpAlaCysSerIle 195
   :::::::::::||||| ||::: ||
451 TTAAACGTGCTGTGCCAGCTGTCTGGGAAAGGCCCGCTCACTGCCTGTGA 500
   :::::::::::||||| ||::: ||
195 ePheAspLeuLysAsnSerLeuProAspAlaLeuLeuProAsnLeuLeu 212
   :::::::::::||||| ||::: ||
501 CTTTGACCTCCGACGCTGCAGCCTGCACAAGCGGCTAGAAAAACCTCCTGC 550
   :::::::::::||||| ||::: ||
212 sPArgThrProAsnGluIleAspArgGlnAsnAspAspGlnArgLys 228
   :::::::::::||||| ||::: ||
551 AGCAAGTGAAGTCCGAGAGCTTTGAGAAGCAGAAAGAGAGGCCCGGAGG 600
   :::::::::::||||| ||::: ||
229 SerAsnArgHisLysGluLeuPheAlaLeuHisProSerProAspGluGlu 245
   :::::::::::||||| ||::: ||
601 ACCAACAGGACGAGCCGAGCTTTGCCCTTACCCTCAGTGAGCAGAGA 650
   :::::::::::||||| ||::: ||
245 uGluProIleGluArgLeuSerValProAspIleProLysGluHisPheG 262
   :::::::::::||||| ||::: ||
651 GGATGCTGTGGAATACGTCCAGTACCAAGATGTCCCAAGGAACACCTGG 700
   :::::::::::||||| ||::: ||
262 LysGlnArgLeuLeuValLysCysLeuSerLeuLysPheGluIleGluIle 278
   :::::::::::||||| ||::: ||
701 GCACAGAAATATTGGTCAAGTTGCTGACCTTGAAGTTGAGATTGAATTT 750
   :::::::::::||||| ||::: ||
279 GluProIlePheAlaSerLeuAlaLeuTyrAspValLysGluLysLysLys 295
   :::::::::::||||| ||::: ||
751 GAGCCCTGTTGGCAGCAATTGCCCTCTACGATGTTAAAGAAAGAAAAA 800
   :::::::::::||||| ||::: ||
295 sIleSerGluAsnPheTyrPheAspLeuAsnSerGluGlnMetLysGlyL 312
   :::::::::::||||| ||::: ||
801 GATCTCAGAAAAATTTTACAGTGTGACCTGAACCTGACACAGTTCAAGGAT 850
   :::::::::::||||| ||::: ||
```

```
312 euleuArgProHisValProProAlaAlaIleThrThrLeuAlaArgSer 328
   :::::::::::||||| ||::: ||
851 TTCTGCGAGCTCACACGCCCTTCAGTGGCCGCAATCAAGTCAGCGAGATCT 900
   :::::::::::||||| ||::: ||
329 AlaIlePheSerIleThrTyrProSerGlnAspValPheLeuValIleLys 345
   :::::::::::||||| ||::: ||
901 GCAGTCTTCTCAGTCACCTACCCGTCCTCAGACATCTACCTGTAAGTCAA 950
   :::::::::::||||| ||::: ||
345 sLeuGluLysValLeuGlnGlnGlyAspIleGlyGluCysAlaGluProT 362
   :::::::::::||||| ||::: ||
951 GATTGAAAAAGTCCCTGCAGCAGCGAGATATTGGAGACTGTGCAGAGCCCT 1000
   :::::::::::||||| ||::: ||
362 yrMetIlePheLysGluAlaAspAlaThrLysAsnLysGluLysLeuGlu 378
   :::::::::::||||| ||::: ||
1001 ACACGGTTATCAAGAAGATGATGGTGGAAGAGTTAAAGAAAGATTGAA 1050
   :::::::::::||||| ||::: ||
379 LysLeuLysSerGlnAlaAspGlnPheCysGlnArgLeuGlyLysTyrAr 395
   :::::::::::||||| ||::: ||
1051 AAACTTAAACTCCAAAGCTGAATCTTCTGCCAGCGTTGGGAAATACCG 1100
   :::::::::::||||| ||::: ||
395 gmetProPheAlaTrpThrAlaIleHisLeuMetAsnIleValSerSera 412
   :::::::::::||||| ||::: ||
1101 GATGCCCTTTGCCCTGGGCACCACATAAGCTTATCAAGCTTCTTTC..AATG 1147
   :::::::::::||||| ||::: ||
412 IagLysLeuGluArgAspSerThrGluValGluIleSerThrGlyGlu 428
   :::::::::::||||| ||::: ||
1148 TCTCCACCCCTTGAGAGGAGGTAACGTGATGTGACCTCTGTGGTGGG... 1194
   :::::::::::||||| ||::: ||
429 ArgLysGlySerTrpSerGluArgAsnSerSerIleValGlyArgAr 445
   :::::::::::||||| ||::: ||
1195 ..AGAACCCAGTGGGTGAACGAGAGACATTTGCC...CAATCTAGAAG 1238
   :::::::::::||||| ||::: ||
445 gSerLeuGluArgThrThrSerGlyAspAspAlaCysAsnLeuThrSerP 462
   :::::::::::||||| ||::: ||
1239 GCTTTGTGAAGAGCCCTCTCTTGGAGGAAATGGGTTGGATCCAACT 1288
   :::::::::::||||| ||::: ||
462 heArgProAlaThrLeuThrValThrAsnPhePheLysGlnGluLysAsp 478
   :::::::::::||||| ||::: ||
1289 TCAAAACCTCCACTCTGAGCGTTAGCAGCTTTTCAACGAGGAAGAGAT 1338
   :::::::::::||||| ||::: ||
479 ArgLeuSerAspGluAspLeuTyrLysPheLeuAlaAspMetArgArgPr 495
   :::::::::::||||| ||::: ||
1339 CGCCTTAGCGATGAAGACTTATTCAAAGTTTTTATGCTGACTACAAAGATC 1388
   :::::::::::||||| ||::: ||
495 oSerSerValLeuArgArgLeuArgProIleThrAlaGlnLeuLysIleA 512
   :::::::::::||||| ||::: ||
1389 ATCATCTCTTACAGAGACGAGTCAATTCACAGCCTTGCTTAAGACTGG 1438
   :::::::::::||||| ||::: ||
512 sPileSerProAlaProGluAsnProHisTyrCysLeuThrProGluLeu 528
   :::::::::::||||| ||::: ||
1439 AGATTCTTACAGCTCCAGAGATCATCAATTGCTGTCTGACTCCTGAATATG 1488
   :::::::::::||||| ||::: ||
529 LeuGlnValLysLeuTyrProAspSerArgValArgProThrArgGluIle 545
   :::::::::::||||| ||::: ||
1489 CTGCCCGTGAACCCCTTCTCGTGAACCCGAGACAGCCCCGCACAAAGAGAT 1538
   :::::::::::||||| ||::: ||
545 eLeuGluPheProAlaArgAspValTyrValProAsnThrThrTyrArgA 562
   :::::::::::||||| ||::: ||
1539 TTTGGAATTTCCAACACGAGAGATATATGCCCTCACACTGTGTACAGAA 1588
   :::::::::::||||| ||::: ||
562 snLeuLeuTyrIleTyrProGlnSerLeuAsnPheAlaAsnArgGlnGly 578
   :::::::::::||||| ||::: ||
1589 ACCTTCTCTATGTCTACCCACAGAGCTGAACCTTGTGAACAACAACTAGCA 1638
   :::::::::::||||| ||::: ||
579 SerAlaArgAsnIleThrValLysValGlnPheMetTyrGlyGluAspPr 595
   :::::::::::||||| ||::: ||
1639 TCAGCCCGGGAACATTACAAATAAAGATCCAGTTTATGTGTGAGAGAAGATGC 1688
   :::::::::::||||| ||::: ||
595 oSerAsnAlaMetProValIlePheGlyLysSerSerCysSerGluPheS 612
   :::::::::::||||| ||::: ||
1689 TAGCAATGGCATGCCGGTCACTTTTGGAAAAATCCAGCGGGCCTGAATTTTC 1738
   :::::::::::||||| ||::: ||
612 erLysGluAlaTyrThrAlaValValTyrHisAsnArgSerProAspPhe 628
   :::::::::::||||| ||::: ||
```


5197 GACAGCATCGTTAACAGGAT.....CATAGAGAAATGTTGGAAACCTA 5240
1832 rpheargValGlyPheTyrGlyThrLysPheGlyAspLeuAspGluGln 1849
5241 CTTCGAGTTGGTTCTTTGGATCCAAATTTGGGATTTGGATGAACAG 5290
1849 lupheValTyrLysGluProAlaIleThrLysLeuAlaGluIleSerHis 1865
5291 AGTTGTCTACAAGAGCCGCAATTACCAAGCTTCCTGAGATCTCACAT 5340
1866 ArgLeuGluGlyPheTyrGlyGluArgPheGlyGluAspValValGluVal 1882
5341 AGACTAGAGGCATTTTATGCTCAATGTTTGGTGCAGAATTTGTGGAAGT 5390
1882 lIleLysAspSerAsnProValAspLysCysLysLeuAspProAsnLysA 1899
5391 GATTAAAGACTCCACTCTCTGTGGACAAACCAAGTTGGATCCTAACCAAG 5440
1899 lATyrIleGlnIleThrTyrValGluProTyrPheAspThrTyrGluMet 1915
5441 CCTACATACAGATCAGATTTTGTGGAGCCCTACTTTGATGAGTATGAGATG 5490
1916 LysAspArgIleThrTyrPheAspLysAsnTyrAsnLeuArgArgPheMe 1932
5491 AAGACACAGGTCACATACTTTGAGAAGAATTTCACCTCCGAGGTTCAT 5540
1932 tYrCysThrProPheThrLeuAspGlyArgAlaHisGlyGluLeuHisG 1949
5541 GTACACCACCCCGTTACCCCTGGAGGGCGCGCTCGGGGAGAGCTGCATG 5590
1949 lueGlnPheLysArgLysThrIleLeuThrThrSerHisAlaPheProTyr 1965
5591 AGCAGTACAGAAAGAACACAGTCTCTGACCACCTATGCACGCCCTTCCCTAC 5640
1966 IleLysThrArgValAsnValThrHisLysGluGluIleIleLeuThrPr 1982
5641 ATCAAGACCAGGATCAGCTCATCCAGAGGAGGAGTTGTTTGGACACC 5690
1982 oIleGluValAlaIleGluAspMetGlnLysLysThrGlnGluLeuAlaP 1999
5691 GATTGAAGTTGCCATTGAAGACATGAAGAAGAAGACCTGCAGTTAGCAG 5740
1999 heAlaThrHisGlnAspProAlaAspProLysMetLeuGlnMetValLeu 2015
5741 TTGCCATTAAACAGAGAGCCGCTGATGCAAAGATGCTTCAGATGCTGCTG 5790
2016 GlnGlySerValGlyThrThrValAsnGlnGlyProLeuGluValAlaGln 2032
5791 CAAGGCTCTGTGGAGCTTCTTAATCAGGGACCACTGGAAGTAGCCCA 5840
2032 nValPheLeuSerGluIleProSerAspProLysLeuPheArgHisHisA 2049
5841 AGTGTTTTGGCTGAATTCCTGCTGATCCAAACCTCTATCGACATCACA 5890
2049 snLysLeuArgLeuCysPheLysAspPheThrLysArgCysGluAspAla 2065
5891 ACAAGTTGAGGTATGCTTAAAGAAATTCATCATGAGATGTGTGAAGCT 5940
2066 LeuArgLysAsnLysSerLeuIleGlyProValGlnLysGluTyrGlnAr 2082
5941 GTAGAGAAAAACAAGCGCTCATCAGGCAGACACAGAGGAATATCAGCA 5990
2082 gGluLeuGlyLys 2086
5991 GGAAGTCAAAAAG 6003

seq_name: /SIDS1/gcgdata/hold-geneseg/genesegn-emb1/NA2001B.DAT:ABLI17736

seq_documentation_block:

ID ABL17736 standard; DNA; 9389 BP.
XX
AC ABL17736;

XX 26-MAR-2002 (first entry)
DT
XX Drosophila melanogaster genomic polynucleotide SEQ ID NO 4681.
DE
XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ds.
XX Drosophila melanogaster.
OS
XX WO200171042-A2.
PN
XX 27-SEP-2001.
PD
XX 23-MAR-2001; 2001WO-US09231.
PF
XX 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
PR
XX (PEKE) PE CORP NY.
PA
XX Venter JC, Adams M, Li PWD, Myers EW;
PI
XX WPI; 2001-656860/75.
DR
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
XX Claim 1; SEQ ID NO 4681; 21pp + Sequence Listing; English.
PS
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABLI6176-ABLI30511), expressed DNA
CC sequences (ABLI01840-ABLI6175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 9389 BP; 2710 A; 1842 C; 1849 G; 2988 T; 0 other;

alignment_scores:
Quality: 4690.00 Length: 2340
Ratio: 2.886 Gaps: 58
Percent Similarity: 69.444 Percent Identity: 44.188

alignment_block:
US-09-737-246-2 x ABLI17736/rev ..

Align seg 1/1 to reverse of: ABLI17736 from: 1 to: 9389

4 ArgArgAlaPheAlaGlnLysIleSer..... 12
8277 CAGCGCTCGTTTGTGCAAAAAGTTTCCAAGTAAGTTTTCACATAAAGAA 8228
12 12
8227 ATAGTTTATTTTAAACAAGAACGCCCTTAATGCTGTATACATATGTA 8178
13ArgThrValAlaAlaGluValArgLys 21
8177 CATATATCAATTGTGAATCCCTAGGACAGTCGGCTGCCGATGTACGAAAA 8128
22 GlnIleSerGlyGlnTyrSerGlySerProGlnLeuLeuLysAsnLeuAs 38
8127 AATGTTTCTGCTTGCATCTATCAAAAGCAATG.....GACCCATC 8087
38 nIleValGlyAsnIleSer.HisHisThrThrValProLeuThr..... 52

568 oGlnSerLeuAsnPhelaIaasnaTrGInglySerAlaArgAsnIleThrV 585
:::|||||:::||||| ||||| ||||| ||||| :::
6290 AAGGAGCTAAATTTTCTTCGGCGTGACAGGTCTGCTCGAACAATTGCCG 6241
585 allysValGlnPheMetYrClgLuAsProSerAsnaIaMetProVal 601
|||||:::||||| ||||| |||||:::|||||:::
6240 TTCGCGTTCAACTAATGCGTGGAGAAACGCCAAAAGATGGCGTCAATGCC 6191
602 IlePheGlyLysSerSerysserGluPheSerLysGluAlaTyrrAl 618
|||||:::|||||:::|||||:::|||||:::|||||
6190 ATCTACGGCAGAAGTCGTTCTGCCAAATTTTCTACTGAAGCATTTACACC 6141
618 aValValTyrrHisAsnaArgSerProAspPheHisGluGluIleLysVal 635
||||| |||||:::|||||:::|||||:::|||||:::
6140 TGTCAATTATCATTAACAATAATGCCCATCATTTCTACGATGAATAAAAAATTG 6091
635 ysLeuProAlaThrLeuThraSPHISHisHisLysLeuPheThrPheTyrr 651
|||||:::|||||:::|||||:::|||||:::|||||
6090 CGCTGCCTGCATCAATTAAGCAGCATCACCATTTATTGTTCCACCATTTTAC 6041
652 HisValserCySGInGlnIlysgln.....AsnThrProLeuGluTh 665
|||||:::|||||:::|||||:::|||||:::
6040 CATGTTTCATGTCAAAGAAACACACAGACTTCGACGCCGTACAGTAGAGAC 5991
665 rProValGlyTyrrThrTriLePrometLeuGlnAsnGlyArgLeuLysT 682
|||||:::|||||:::|||||:::|||||:::|||||
5990 TCCTATCGGCTACACATGGTTGCCCTTACTGGAAGATGGAACAATTAAAGT 5941
682 hrGlyGlnPheCysLeuProValserLeuGluLysProProglnAlaTyrr 698
|||||:::||||| |||||:::|||||:::|||||:::
5940 TTGGGGAATTAACTTCTCGTTATGTTAGTAATCACC GCCGGAATAATTAC 5891
699 serValleuSerProgluValProleuProglumetystrValaspas 715
|||||:::|||||:::|||||:::|||||:::|||||
5890 TCATTATTATACCACCGAATGTTCACTGCCCTGGAATTAAATGGCTGGACAA 5841
715 nhIsLySGlyValPheaSnValGluValAlaValaserSerIleHist 732
|||||:::|||||:::|||||:::|||||:::|||||
5840 TCACAGAGCCGCTTTCTATTAAATGTAGAAGCAGTCAACCAATTACATA 5791
732 hrGlnAspProTyrrLeuAsplysPhePheAlaLeuValasnAlaLeuasp 748
|||||:::|||||:::|||||:::|||||:::|||||
5790 CTTTGATTTCTTCCCTCGATCGATTTCTTTAATTGGCAATATCTAGAC 5741
749 GluHisLeuPheProValArgIleGlyAspmetArgIlemetGluasnas 765
:::|||||:::|||||:::|||||:::|||||
5740 ACTCGAAATATACCTTCCCATATTGGC.....GAGAAATAA 5706
765 nleuGluAsngluLeuLysSerSerIleseralAlaLeuAsnSerSerglnl 782
|||||:::|||||:::|||||:::|||||:::|||||
5705 TATAGAAACAGAATTAAAGAAATGCTTGGCTGATATCGAATATGCAAAATC 5656
782 euGluProValValArgPheLeuHisLeuLeuLeuAsplysLeuIleLeu 798
|||||:::|||||:::|||||:::|||||:::|||||
5655 GTGAGCCATTAGTGAGGCATCTACCATTTAGTCTTGATTAAGCTTATCGAA 5606
799 leuValIleargProProValIlealaglyGlnIlevalasnlleuglyl 815
|||||:::|||||:::|||||:::|||||:::|||||
5605 TTGTAGTTGTACACCACAAGTTGGTGACAAAGCAATGCTCTGGGATC 5556
815 naIaserPheGluAlametalaserIleIeAsnArgLeuHisLysAsnl 832
:::|||||:::|||||:::|||||:::|||||
5555 TACGCTTTTCGAAGTCTCTTCTTGGTATCATCTTTGTTATCAATACTTA 5506
832 eucGluGLyAsnhisaspGlnhisGlyArgAsnSerLeuLeualaSerTyrr 848
:::|||||:::|||||:::|||||:::|||||:::|||||
5505 ACGACGATCAATATGACCAATACGGTCCGCCAAAGTTTGCCTATCAACATAT 5456
849 IlehIstYrrValPheArgLeuProAsnthTyrrProAsnSerSerSerPr 865
|||||:::|||||:::|||||:::|||||:::|||||
5455 GTGCACCTTTCATGTAAATTCGCCACCCCTTTCAA..... 5420

865	oGlyProGlyGlyLeuGlyGlySerValHisTyrAlaThrMetAlaArgS	882
5419AGCAAAACAGCGACTAACATGCAGTCGCA	5392
882	erAlaValArgProAlaSerLeuAsnLeuAsnArgSerArgSerLeuSer	898
5391	GCACAACAGAA....GAC TTGCAATTAAAGCAATCGTACACCATATAT	5348
899	AsnSerAsnProAspIleSerGlyThrProThrSerProAspAspGluVa	915
5347	GATAAT.....	5342
915	IArgSerIleIleGlySerLysGlyLeuAspArgSerAsnSerTrpValA	932
5341	...GTTTAGCAAGTGGCCGAAGCTTGATGCGCAAAAGTAA.....A	5303
932	snThrGlyGlyProLysAlaAlaProTrpGlySerAsnProSerProSer	948
5302	ATACATATTTGTAAAGCT.....	5283
949	AlaGluSerThrGlnAlaMetAspArgSerCysAsnArgMetSerSerHi	965
5282TTAAATATAACCTTTTCCT	5263
965	sThrGluThrSer...SerPheLeuGlnThrLeuThrGlyArgLeuProT	981
5262	TTCAAGAGCTGTCATAGATGTCCTTCACTCGATGGCTGCACGGGATGCC	5213
981	hrlLysLysLeuPheHisGluGluLeuAlaLeuGlnTrpValValCysSer	997
5212	AAGTGGCGCTGCTTCACGAAGAATTAAGCCTTGCAATGGTGTTCGCAAGT	5163
998	GlySerValArgGluSerAlaLeuGlnGlnAlaTrpPhePheGluLe	1014
5162	GGAAGGCGAGCTGACTTAGCCATGTCAAATTCGTGTTTATTTGAGCT	5113
1014	uMetValLysSerMetValHisIleLeuTyrPheAsnAspLysLeuGluA	1031
5112	TATGTCAAATCTATGATGTAACATTTGCATGTGTCAAACTTTAAATG	5063
1031	lProArgLysSerArgPheProGluArgPheMetAspIleAlaAla	1047
5062	GTCCCGGCAACATCGATTTCCACATCAGTTTATGATGATCTATCCACC	5013
1048	LeuValSerThrIleAlaSerAspIleValSerArgPheGlnLysAspTh	1064
5012	CTGGTGCACTTAGTTACACAACAAAGTGTGGCTACACAGCAACGAACC	4963
1064	rGluMetValGluArgLeuAsnThrSerLeuAlaPhePheLeuAsnAspL	1081
4962	TAAATTAGCGCAGTCATTAATGCAAGTTTAGCTTTTATATTTGACA	4913
1081	euleuSerValMetAspArgGlyPheValPheSerLeuIleLysSerCys	1097
4912	TACTAAGCATCATGGATGGGAGTTTGTTCGGCCTCATCAAAACGTAC	4863
1098	TyrLysGlnValSerSerLysLeuTyrSerLeuProAsnProSerValIle	1114
4862	ACAAAGGTTCTGATTTCTAAAAATGCTTCAATACCGGAT.....CT	4822
1114	uValSerLeuArgLeuAspPheLeuArgIleIleCysSerHisGluHisT	1131
4821	AATGAACATATAAAATAGACTTTTGGAGAATTGCTGTAGTCATGAGCAAT	4772
1131	yValThrLeuAsnLeuPro.....CysSerLeuLeuThrPro	1143
4771	TTGTAGCCTTAAACTTGGCGTTTGGTACTTCAATACACAATGGTAAACAGCG	4722
1144	ProAlaSerProSerProSerValSerSerAlaThrSerGlnSerSergl	1160
4721	CCATGCAGCCCCACGCCAAGCACACAACGTGCAGCAACGCCAAACTTCTTG	4672
1160	yPheSerThrAsnVal.....	1165

```
1111 : : : : :
4671 CGTAGTAAAAAACAATACACTTAAATTTTACTTATTGTGGCTTCTT 4622
1166 ....Glnaspgln..LysIleAlaAsnMetPheGluLeuSerValProp 1180
      : : : : : : : : : : : : : : : : : : : : : : : :
4621 TTAAAAAGGATCAGTACAGAGAGAGCCCTTCAGCGCCGACCTTAGTCAGAAAT 4572
1180 hearGgInGlnHisTyrLeuAlaGlyLeuValLeuThrGluLeuAlaVal 1196
      ||||| : : : : : : : : : : : : : : : : : : : : : : : :
4571 TCCTCCAGCAACATTTTCTAGTCGACCTGTCTTGAGCGACTTAGACGA 4522
1197 IleLeuAsp..ProAspAlaGluGlyLeuPheGlyLeuHisTyrSlySva 1212
      : : : : : : : : : : : : : : : : : : : : : : : :
4521 GTAATGGAGTGGCCAAACCCCA.....TTACATGMAAAGC 4484
1212 IleAsnMetValHisAsnLeuLeuSerSerHisAspSerAspProArgT 1229
      : : : : : : : : : : : : : : : : : : : : : : : :
4483 AATAAGGTGATTCGCAACCTGATGACATCTCACGATTTGGACGCGCGGT 4434
1229 ySerAspProGlnIleLysAlaArgValAlaMetLeuTyrLeuProLeu 1245
      ||||| : : : : : : : : : : : : : : : : : : : : : : : :
4433 ATAGCGAAACTGATGCACGCTGCAGAGGGTGCCTTATATATACCGCTA 4384
1246 IleGlyIleIleMetGluThrValProGlnLeuTyrAspPhe..... 1259
      : : : : : : : : : : : : : : : : : : : : : : : :
4383 CTTTCTATAGTAGATGGATTCATTCCTCAGTTGCATCAGCACGCGCTTGA 4334
1260 .ThrGluThrHisAsnGlnArgGlyArgProIleCysIleAlaThrAspa 1276
      : : : : : : : : : : : : : : : : : : : : : : : :
4333 GCAGGATCGTCTGCAGCAAAATCGGTCAA.....CTTGAAG 4299
1276 sPTyrGluSerGlu.....SerGlySerMetIleSerGlnThr 1288
      ||||| : : : : : : : : : : : : : : : : : : : : : : : :
4298 ACTATCAAGGGCCACATCAAACTATATACAGCATCAACCATAAATCCGGAA 4249
1289 ValAlaMetAlaIleAlaGlyThrSerValProGlnLeuThrArgProG 1305
      ||||| : : : : : : : : : : : : : : : : : : : : : : : :
4248 GTGGCATTTGGCATATCGGGA.....AGTCGTCCCTA 4217
1305 ySerPheLeuLeuThrSerThrSerGlyArgGlnHisThrThrPheSera 1322
      ||||| : : : : : : : : : : : : : : : : : : : : : : : :
4216 CTCCTACTTG.....AATGACACAGTTAAACAAATTGCCCTCAGCT 4173
1322 IagLuserSerArgSerLeuLeuIleCysLeuLeuTrpValLeuLysAsn 1338
      : : : : : : : : : : : : : : : : : : : : : : : :
4172 CGGAGAATACCCGCAATTATTGTGTGTTTCTGTGGTGCTATAAAAAAC 4123
1339 AlaAspGluThrValLeuGlnLysTrpPheThrAspLeuSerValLeuG 1355
      : : : : : : : : : : : : : : : : : : : : : : : :
4122 TTGGAGCGCAACGTTTATATCGCTGGCTTATGATTTAAGTCCGATAG 4073
1355 nLeuAsnArgLeuLeuAspLeuLeuTyrLeuCysValSerCysPheGluTr 1372
      : : : : : : : : : : : : : : : : : : : : : : : :
4072 AGTGCACCAAAATGCTTCAAGGTGTAAACGTTTGCTTGAAAAACTTTCGAGT 4023
1372 yRLysGlyLysLys.....ValPheGluArgMetAsnSerLeuThr 1385
      ||||| : : : : : : : : : : : : : : : : : : : : : : : :
4022 ACACAGGACAAAAAAGCGTGGCCACCCCTTAAACGAACGATACACAAAGT 3973
1386 PheLysLys.....SerLysAspMetArgAlaLysLeuGluAlaIle 1400
      ||||| : : : : : : : : : : : : : : : : : : : : : : : :
3972 TTTAGGAAAAACTGGATCTACTGATGTTAAGGAAAAAGCTGGAGAAGATGTAT 3923
1400 eleuGlySerIleGlyAlaArgGlnGluMetValArgArgSerArgGlyG 1417
      ||||| : : : : : : : : : : : : : : : : : : : : : : : :
3922 TAGAGGCACAAATTCAGCACGTTATGATTTAATAATCGTGAAAAAGATC 3873
1417 InLeuGluArgSerProSerGlySerAlaPheGlySerGlnGluAsnLeu 1433
      : : : : : : : : : : : : : : : : : : : : : : : :
3872 GT.....AATTCAAACAGAAAAAATTC 3853
1434 ArgTrpArgLysAspMetThrHisTrpArg...GlnAsnThrGluLysLe 1449
      ||||| : : : : : : : : : : : : : : : : : : : : : : : :
```

```
3852 CGATGGAGGAAGACAGATGCCATATCGCTCTCAATATATCGGATGCTGT 3803
1449 uAspLysSerArgAlaGluIleGlnHisGluAlaLeuIleAspGlyAsnL 1466
      : : : : : : : : : : : : : : : : : : : : : : : :
3802 GGGAAAAAGTAGACATGAACCTTGAATTAAGTCATTTCAATCGAAGGCTCTT 3753
1466 euAlaThrGluAlaAsnLeuIleIleLeuAspThrLeuGluIleValVal 1482
      ||||| : : : : : : : : : : : : : : : : : : : : : : : :
3752 TGGCAACTGAAGTTGCACCTGTACTTCTCGATACCCCTTGAATAATTTGTT 3703
1483 GlnThrValSerValThrGluSerLysGlySerIleLeuGlyGlyValLe 1499
      : : : : : : : : : : : : : : : : : : : : : : : :
3702 CAT.....GCTGCTGCCAACCTTTATCACAATCTTCTTGAACGTGTGCT 3659
1499 uLysValLeuLeuHisSerMetAlaCysAsnGlnSerAlaValTyrLeuG 1516
      ||||| : : : : : : : : : : : : : : : : : : : : : : : :
3658 AAAGGTGCTTCTTCATTTCTTTGTCCCGGAATCAATCTGTTTGGCCCTGC 3609
1516 InHisCysPheAlaThrGlnArgAlaLeuValSerLysPheProGluLeu 1532
      ||||| : : : : : : : : : : : : : : : : : : : : : : : :
3608 AAAACTTGTGTGCTCCCAACGCGCGTTAATCTTTAAATTCCAAACCTTA 3559
1533 LeuPheGluGluGluThrGluGlnCysAlaAspLeuCysLeuArgLeuLe 1549
      ||||| : : : : : : : : : : : : : : : : : : : : : : : :
3558 TTATTTGATGACGAGACCATATATGTGCGGATTTGTCTAATTTTACT 3509
1549 uArgHisCysSerSerSerIleGlyThrIleArgSerHisProSerAlaAs 1566
      : : : : : : : : : : : : : : : : : : : : : : : :
3508 GAAACATTTGTGGGTCCCTATTTGCCAGGAATACGATCACAGGACGTGCAT 3459
1566 erLeuTyrLeuLeuMetLysGlnAsnPheGluIleGlyAsn..... 1579
      ||||| : : : : : : : : : : : : : : : : : : : : : : : :
3458 CACTATATTACTAATGAGACAAAAATTTTGAATTTGGAATGTAAAGTTT 3409
1579 ..... 1579
3408 ATTTTATTGTTAGGTTTATTAACTAAATTAATGCTCACTTATTCGTAT 3359
1580 .Asn..PheAlaArgValLysMetGlnValProMetSerLeuSerSerLeu 1595
      : : : : : : : : : : : : : : : : : : : : : : : :
3358 AGAACTTTGCCCGAGTTAAGATGCAGAGTGACGATGTCTTTAAGCTCCCTA 3309
1596 ValGlyThrSerGlnAsnPheAsnGlnGluPheLeuArgArgSerLeuLy 1612
      ||||| : : : : : : : : : : : : : : : : : : : : : : : :
3308 GTTGGAAACAAGTTCCGTTCTTTAGTAGCAATCCCTTAACGCCGTGCTTA 3259
1612 sThrIleLeuThrTyrAlaGluGluAspLeuGluLeuArgGluThrThrP 1629
      ||||| : : : : : : : : : : : : : : : : : : : : : : : :
3258 AACCGTCTCGTTATGTGTAATCCGACTCCGACTGCAGGATACTTCTT 3209
1629 heProAspGlnValGlnAspLeuValPheAsnLeuHisMetIleLeuSer 1645
      ||||| : : : : : : : : : : : : : : : : : : : : : : : :
3208 TTCCTGAACAAGTACAAAGATTTCCTTTTAAATGTCATATGATCCTGTCA 3159
1646 AspThrValLysMetLysGluHisGlnGluAspProGluMetLeuIleAs 1662
      ||||| : : : : : : : : : : : : : : : : : : : : : : : :
3158 GATACTGTTTAAATGAAGAGATATCAGGAAGACCCAGAAATGTTGCTTGA 3109
1662 PleuMetTyrArgIleAlaLysGlyTyrGlnThrSerProGlu...ArgL 1678
      ||||| : : : : : : : : : : : : : : : : : : : : : : : :
3108 CCTCATGAATCGTATTCGCAAGGATATACCAAAATAACCTGATCTACGAC 3059
1678 euThrTrpLeuGlnAsnMetAlaGlyLysHisSerGluArgSerAsnHis 1694
      ||||| : : : : : : : : : : : : : : : : : : : : : : : :
3058 TGACTTGGTTGGAATAATATGGCTTAAAAAACACCCGAGCGAACAATCAC 3009
1695 AlaGluAlaAlaGlnCysLeuValHisSerAlaAlaLeuValAlaGluTy 1711
      : : : : : : : : : : : : : : : : : : : : : : : :
3008 ACGGAAGCAGCCATGTGTATGTACATGCTGCTTCTTTAGTTTCTGAATA 2959
1711 rLeuSerMetLeuGluAspArgLysTyrLeuProValGlyCysValThrP 1728
      ||||| : : : : : : : : : : : : : : : : : : : : : : : :
2958 TCTTAGCATGTGTGAGTCACAAAAACATTTGCTGTGGAGCTGTAGATT 2909
```

1728 heGlnAsnIleSerSerAsnValIleuGluIleuSerAlaValSerAsp 1744
|||||:||||| |||:||||| ||||| ||||| ||||| |||||
2908 TTCAACGAATTTCTCCCAACACACTAATGAGTCGCCGTATCGGATGAT 2859
1745 ValValSerProAspGluGluGlyIleCysSerGlyLysThrGlu 1761
|||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
2858 GTGCTAAGTCCCCGGGAAGATGATCTGCCCTAGGAATCATTTCACTGA 2809
1761 uSerGlyLeuValGlyLeuLeuGluGlnAlaAlaSerPheSerMet 1778
|:||||| :|:|||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
2808 AACTGGGTTGAAGGCCCTTGCTGGAAGAAGCCTCCAATTCTTTCAAGTTG 2759
1778 laGlyMetThrGluAlaValAsnGluValTyrLysValIleProIle 1794
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2758 CTGCATGTATGAACAAATGAAGAGTGTACAAAATTCATAATACCATA 2709
1795 HisGluAlaAsnArgAspAlaLysLysLeuSerThrIleHisGlyLysLe 1811
|:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2708 TGCAGAGCTAACAGAGATTTTCAAAAGCTAAGCAAAGTTCAATGGCAAAATT 2659
1811 uGlnGluAlaPheSer..... 1816
|:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2658 GCAGAGAGCATTTAATCGAATATCCCACTACAGGTACAAATATTGTGTA 2609
1817LysIle.ValHisGlnSerThrGlyTyrPgluArgMe 1828
2608 AATTTTACCACACGGAATAATATATACATATTATAACAGGGTAAGAGAGT 2559
1828 tPheGlyThrTyrPheArgValGlyPheTyrGlyThrLysPheGlyAspL 1845
:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2558 TTTTGGACATACTTTCGTGTGGCTTCTATGGCGGAATAATTGGGACT 2509
1845 euAspGluGlnGluPheValTyrLysGluProAlaIleThrLysLeuAla 1861
|||||:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2508 TGGATCAGCAGGAATTCATTTATTAAGAGCCAACATTTGACGAGTTGCC 2459
1862 GluIleSerHisArgLeu.Glu..... 1868
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2458 GAAATATTTAGTCGGCTTCAGGTATATATTGCAAAATTGGAATAATAGAA 2409
1869GlyPheTyrGlyGluArgPhe 1875
2408 CTAATCAATTTTGTTCACACATACGTTAGAACTTTTACACTGAAACGATTTC 2359
1876 GlyGluAspValValGluValIleLysAspSerAsnProValAspLysCy 1892
||| ||| |||:||||| ||||| ||||| ||||| ||||| ||||| |||||
2358 GGACCGGACTCTGTGCATATCATTAAGATTTCCAATACCGTTGATATTAA 2309
1892 sLysLeuAspProAsnLysAlaTyrIleGlnIleThrTyrValGluProT 1909
:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2308 TAGCTTGGATCCCGATTAAGGCTTACATTTCAAAATTACTTATGTGAACCCCT 2259
1909 yrPheAspThrTyrGluMetLysAspArgIleThrTyrPheAspLysAsn 1925
|||||:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2258 ACTTTGAACATATGAATGCGTCATCGTGAGACATACTTTGACCGGAAT 2209
1926 TyrAsnLeu..... 1928
:|||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
2208 TTCATATATAGTATGATATGAATTAACAGATAATTAAATCGAAATTTTA 2159
1929Arg.ArgPheMetTyrCys 1934
||| |||||:|||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
2158 AAATTAATAATTTTAACATTTTGGTTTATGCAAGACGTTTATATATATGCC 2109
1935 ThrProPheThrLeuAspGlyArgAlaHisGlyGluLeuHisGluGlnPh 1951
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2108 ACGCCTTTTACTAAAAATGGAAGGCACATGGCGAACTAAATGAGCAGTG 2059
1951 elysArgLysThrIleLeuThrThrSerHisAlaPheProTyrIleLysT 1968
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2058 TAAACGGAAACGATTTTGACACACGGCAATCATCTTCCATACGTAAAAA 2009

1968 hrArgValAsnValThrHisLysGluGluIleIleLeuThrProIleGlu 1984
|||||:| ||| :|||:||||| ||||| ||||| ||||| |||||
2008 CTCGCATTTATGTTAATTAGCCGACAGCAAAATAGTCTCGAACCGATTGAA 1959
1985 ValAlaIleGlu..... 1988
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1958 GTGCAATTAAGTAAGTATATGTTGTAATAATCCAAAAGTTAAATAATT 1909
1989AspMetGlnLysThrGlnGluLeuAlaPheA 2000
1908 AAACGTTATTCCTAGATATATCAAAAAAAACATTGGAATTCGACAGCTG 1859
2000 laThrHisGlnAspProAlaAspProLysMetLeuGlnMetValIleuGln 2016
|||||:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1858 CCACGAATCAAGAGCCACAGACACCCCAAAATATTGCAATGCTACTGCAG 1809
2017 GlySerValGlyThrThrValAsnGlnGlyProLeuGluValAlaGlnVa 2033
||| :|:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1808 GGATGTATTGGAACCACTGTTAACCAAGACCGATGGAATGGCGAGTGT 1759
2033 lPheLeuSerGluIleProSerAspProLysLeu..PheArgHisHisA 2049
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1758 GTTCCTTCCAAATTTATCCGACGGAACAACTGTACCGACAAACACCAA 1709
2049 snLysLeuArgLeuCysPheLysAspPheThrLysArgCysGluAspAla 2065
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1708 ACAAACTTCGTTGTGCTTTCGCGAGTTTCAAAAGCGTTGTGCTGATGCT 1659
2066 leuArgLysAsnLysSerLeuIleGlyProValGlnLysGluTyrGlnAr 2082
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1658 TTGAAAAAGAAATCGCAATCTAATACTTTCAAGATCAAAAAGATTACCAACG 1609
2082 gGluLeu 2084
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1608 AGAACTG 1602

seq_name: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA2001A.DAT:AAH41952

seq_documentation_block:
ID AAH41952 standard; DNA; 3173 BP.
XX
AC AAH41952;
XX
DT 30-AUG-2001 (first entry)
XX
DE CLASP-3 nucleotide fragment.
XX
KW Human; cadherin-like asymmetry protein; CLASP; CLASP-3; immune response;
KW cell surface molecule; transmembrane protein; immunosuppressive; vaccine;
KW antiinflammatory; antiarthritic; antianaemic; dermatological; uropathic;
KW ophthalmological; antirheumatic; nephrotropic; antithyroid; antidiabetic;
KW neuroprotective; antiasthmatic; antibacterial; antisense therapy;
KW gene therapy; chromosome 1p31.1; ds.
XX
OS Unidentified.
XX
XX WO200142297-A2.
PN
XX
PD 14-JUN-2001.
XX
PF 13-DEC-2000; 2000WO-US34171.
XX
PR 13-DEC-1999; 99US-0170453.
PR 14-JAN-2000; 2000US-0176195.
PR 14-FEB-2000; 2000US-0182296.
PR 11-APR-2000; 2000US-0196267.
PR 11-APR-2000; 2000US-0196460.
PR 11-APR-2000; 2000US-0196527.
PR 11-APR-2000; 2000US-0196528.
PR 11-APR-2000; 2000US-0547276.
PR 13-OCT-2000; 2000US-0240503.
PR 13-OCT-2000; 2000US-0240508.
XX

PA (ARBO-) ARBOR VITA CORP.
XX
PI Lu P, Garman JD, Candia AF;
XX
DR WPI; 2001-375003/39.
XX
PT Novel isolated cadherin-like asymmetry protein (CLASP) -3, useful for
PT useful in preventing or treating a CLASP-3-mediated disease e.g.
PT autoimmune disease -
XX
PS Disclosure; Fig 9A; 189pp; English.
XX
CC The present invention describes an isolated polypeptide (I) comprising a
CC nucleotide (nt) sequence (Sia) that has at least 90% identity to the
CC sequence given in AAH41934 and is immunologically cross-reactive with
CC the derived amino acid (aa) sequence (S1b) given in AAB99495 or shares a
CC biological function with native CLASP-3. (I) has immunosuppressive,
CC antiinflammatory, antiarthritic, antianaemic, dermatological, uropathic,
CC ophthalmological, antirheumatic, nephrotropic, antithyroid, antidiabetic,
CC neuroprotective, antiasthmatic and antibacterial activities, and can be
CC used in antisense therapy, vaccine production and gene therapy. CLASP-3
CC related sequences can be used in preventing or treating a CLASP-3
CC mediated disease, preferably an autoimmune disease by inhibiting an
CC immune response. The autoimmune disease is caused or exacerbated by
CC increased activity of T helper 1 lymphocytes (TH1). AAH41893 to AAH41953
CC and AAB99491 to AAB99507 represent sequences which are used in the
CC exemplification of the present invention. CLASP-3 is localised in the
CC chromosome location 1p31.1.
XX
SQ Sequence 3173 BP; 981 A; 624 C; 688 G; 880 T; 0 other;

alignment_scores:
Quality: 4600.50 Length: 904
Ratio: 5.106 Gaps: 1
Percent Similarity: 99.668 Percent Identity: 99.115

alignment_block:
US-09-737-246-2 x AAH41952 ..

Align seg 1/1 to: AAH41952 from: 1 to: 3173

1188 GlyLeuValLeuThrGluLeuAlaValIleLeuAspProAspAlaGlu1 1204
|||||
4 GGACTTGTGTTTCAGAGCTGGCTGTCATTTTAGACCCGTGATGCTGAAG 53
1204 yLeuphegLyLeuHisLysLysValIleasnMetValHisasnLeu1221
|||||
54 ACTGTTTGATGTCATAAGAAAGTCATCATATGTGTACACAAATTACTCT 103
1221 erSerHisAspSerAspProArgTyrSerAspProGlnIleLysAlaArg 1237
|||||
104 CCAGTCACGACTCAGACCCGCGGTACTCTGACCCCTCAGATAAAGCTCGA 153
1238 ValAlaMetLeuTyrLeuProLeuIleGlyIleIleMetGluThrValPr 1254
|||||
154 GTGGCCATGTTGATCTACCTCTGATGTGTTATTCATGGAACCTGTACC 203
1254 oGlnLeuTyrAspPheThrGluThrHisAsnGlnArgGlyArgProIleC 1271
|||||
204 TCAGCTGTATGATTTTACAGAACTCACAATCAACGAGGAAGCAATTT 253
1271 ysIleAlaThrAspAspTyrGluSerGluSerGlySerMetIleSerGln 1287
|||||
254 GTATAGCCACTGATGATTATGAAGTGAGAGCGGAAGTATGATAAGCCAG 303
1288 ThrValAlaMetAlaIleAlaGlyThrSerValProGlnLeuThrArgPr 1304
|||||
304 ACCGTTGCCAATGCAATCGCAGGACATCGGTCCCTCAACATAACAAGCC 353
1304 oGlySerPheLeuLeuThrSerThrSerGlyArgGlnHisThrThrPheS 1321
|||||
354 TGGCAGTTTCTCTCTCAGTCACAGAGTGGCAGGCAACACACTACTTTT 403

1321 eAlaGluSerSerArgSerLeuLeuIleCysLeuLeuTrpValLeuLys 1337
|||||
404 CAGCAGAAATCAAGTCGAAGCCTTTTGATCTGTCTACTTTGGGTTCTCAAA 453
1338 AsnAlaAspGluThrValLeuGlnLysTrpPheThrAspLeuSerValLe 1354
|||||
454 AATTCAGATGAACAGTCTTACAGAAAGTGGTTACAGATCTCTCAGTCTT 503
1354 uGlnLeuAsnArgLeuLeuAspLeuLeuTyrLeuCysValSerCysPheG 1371
|||||
504 GCAGCTAAACCGGCTATAGATCTGCTTTATCTCTGTGTCTTGTGCTTTG 553
1371 LuTyrLysGlyLysLysValPheGluArgMetAsnSerLeuThrPheLys 1387
|||||
554 AGTATAAAGGGAAGGAAAGTCTTGAACGAATGAATAGCTTGACCTTTAG 603
1388 LysSerLysAspMetArgAlaLysLeuGluGlnAlaIleLeuGlySerI 1404
|||||
604 AATCAAAAGACTAGAGAGCAAAAGCTTGAGAGAGCTATTCTGGGAGCAT 653
1404 eGlyAlaArgGlnGluMetValArgArgSerArgGlyGlnLeuGluArgS 1421
|||||
654 AGGTGCCAGGCAAGAAATGTTACGGCGAAGCCGAGAGACAGCTCGAGAGA 703
1421 erProSerGlySerAlaPheGlySerGlnGlnLysLeuArgTrpArgLys 1437
|||||
704 GCCCATCTGGAAGTGCTTTGGAAAGTCAAGAAATTTGAGGTGAGAGAA 753
1438 AspMetThrHisTrpArgGlnAsnThrGluLysLeuAspLysSerArgAl 1454
|||||
754 GATATGACTCAGCTGGCGTCAAAACACAGAGAGCTTGACAAATCAAGAGC 803
1454 agLileGluHisGluAlaLeuIleAspGlyAsnLeuAlaThrGluAlaA 1471
|||||
804 AGAGATTGAACACAGACACTGATTGATGGAACCTGGCTACAGAAAGCAA 853
1471 snLeuIleIleLeuAspThrLeuGluIleValValGlnThrValSerVal 1487
:::
854 AGTAATCATTTTATGATACATTAAGATTGTTGTCAGACCGTTTCTGTA 903
1488 ThrGluSerLysGluSerIleLeuGlyGlyValLeuLysValLeuLeuH 1504
|||||
904 ACGGAATCCAAAGAGAGCATTTCTTGGTGAGTGTCTAAAGTGCTACTACA 953
1504 sSerMetAlaCysAsnGlnSerAlaValTyrLeuGlnHisCysPheAlaT 1521
|||||
954 CAGCATGGCCGTGAACCAAGTGAGTTTATCTACAACACACTGTTTGCTA 1003
1521 hrGlnArgAlaLeuValSerLysPheProGluLeuLeuPheGluGlu 1537
|||||
1004 CACAGAGAGCCTTGTTCAAAAGTTCTCCTGAACCTTATTATTGAAGAAGAG 1053
1538 ThrGluGlnCysAlaAspLeuCysLeuArgLeuLeuArgHisCysSerSe 1554
|||||
1054 ACAGAGCAGTGTGCTGATTATGCTCAGGCTTCTCCGACACTGACGACAG 1103
1554 rSerIleGlyThrIleArgSerHisProSerAlaSerLeuTyrLeuLeuM 1571
|||||
1104 TAGCATCGGTACAAATACGGTCAACACCCAGTGCCCTTACCTACTAA 1153
1571 eTArgGlnAsnPheGluIleGlyAsnAsnPheAlaArgValLysMetGln 1587
|||||
1154 TGAGGCAAAACTTTGAGATTGGGAATAACTTTGCCAGGGTTAAATAATCAG 1203
1588 ValPrometSerLeuSerSerLeuValGlyThrSerGlnAsnPheAsnG 1604
|||||
1204 GTACCAATGTCACTATCTCTTGGTGCGACATCTCAGAAATTTTAATGA 1253
1604 uGluPheLeuArgArgSerLeuLysThrIleLeuThrTyrAlaGluGlnA 1621
|||||
1254 AGAATTCTTAAGAGCTTCTCTAAAGACATATATTGACATATGCTGAAGAAG 1303

1621 spleuGluleuArgGluThrThrPheProAspGlnValGlnAspLeuVal 1637
1304 ATCTGGAATTGAGGGAACACATTTCTCTGATCAGGTCCAGGATCTGGTT 1353
1638 PheAsnLeuHisMetIleLeuSerAspThrValLysMetLysGluHisG1 1654
1354 TTCATCTCCATATGATGATCTTCTTGATACTGTGAATAATGAAGAACACCA 1403
1654 ngluAspProGluMetLeuIleAspLeuMetTyrArgIleAlaLysGlyT 1671
1404 GGAGGATCCTGAAATGTGATGATCTAATGTACAGAAATTGCCAAGGCTT 1453
1671 yrglnThrSerProGlu..ArgLeuThrTrpLeuGlnAsnMetAlaGly 1686
1454 ACCAGACCTCTCCAGATCTGCCGATTGACCTGGTTGCAGAACATGGCAGGC 1503
1687 LysHisSerGluArgSerAsnHisAlaGluAlaAlaGlnCysLeuValHi 1703
1504 AAGCACTCAGAACGAAACCAATCATGCTGAAGCTGCACAGTGTCTAGTCCA 1553
1703 sSerAlaAlaLeuValAlaGluTyrLeuSerMetLeuGluAspArgLysT 1720
1554 CTCAGCAGCAGCTTGTGCTGAATATTGAGCATGCTGAGGAGACCGGAAAT 1603
1720 yrlLeuProValGlyCysValThrPheGlnAsnIleSerSerAsnValLeu 1736
1604 ATCTTCCTGTCCCATGTGTACACATTTCCAGAAATATTCTAATGTTTTA 1653
1737 GluGluSerAlaValSerAspAspValValSerProAspGluGluGlyIl 1753
1654 GAAGAATCTGCGGTCTCAGATGATGTGTATCTCCAGATGAAGAAGGTAT 1703
1753 ecYsSerGlyLysTyrPheThrGluSerGlyLeuValGlyLeuLeuGluG 1770
1704 CTGCTCTGGAATAATACTTTACTGAGTCAGGACTGTGGGATTTACTGGAAC 1753
1770 lnaIaIaAlaSerPheSerMetAlaGlyMetTyrGluAlaValAsnGlu 1786
1754 AAGCAGCTGCTTCTCTCTATGGCTGGCATGTATGACACAGTTAATGAA 1803
1787 ValTyrLysValLeuIleProIleHisGluAlaAsnArgAspAlaLysLy 1803
1804 GTTTACAAAGTACTTATTTCTATTCATGAAGCTAATCGGATGCCAAGAA 1853
1803 sLeuSerThrIleHisGlyLysLeuGlnGluAlaPheSerLysIleValH 1820
1854 ACTATCCACAATTCATGTGTAACCTTCAAGAAGCATTCAGCAAAATTGTTTC 1903
1820 lsglnSerThrGlyTyrGluArgMetPheGlyThrTyrPheArgValGly. 1836
1904 ATCAGAGTACTGGCTGGGAGCGGATGTTTGGCACCTATTTTCGTGTGCT 1953
1837 PheTyrGlyThrLysPheGlyAspLeuAspGluGlnGluPheValTyrLy 1853
1954 TTTTATGGAAACCAAGTTCGGGGATTTGGAATGAACAAGAAATTTGTTACAA 2003
1853 sgluProAlaIleThrLysLeuAlaGluIleSerHisArgLeuGluGlyP 1870
2004 GGAGCCTGCAATTAACCAAACTTGCAGAGATATCTCACAGATTGGAGGGAT 2053
1870 heTyrGlyGluArgPheGlyGluAspValValGluValIleLysAspSer 1886
2054 TTTACGGAGAAAGATTTGGAGAGATGTGTTGAAGTAATCAAAAGACTCT 2103
1887 AsnProValAspLysCysLysLeuAspProAsnLysAlaTyrIleGlnIl 1903
2104 AATCCTGTAGACAAGTGTAAATTAGATCCTAACAAGGCATATATTCAGAT 2153
1903 eThrTyrValGluProTyrPheAspThrTyrGluMetLysAspArgIleT 1920
2154 TACCTATGTGAGCCATACTTTGACACATATGAGATGAAGGACAGAATCA 2203
1920 hrTyrPheAspLysAsnTyrAsnLeuArgArgPheMetTyrCysThrPro 1936

2204 CCTATTTCGACAAAAAATTACAAATCTTCGTCGATTCATGTACTGTACACCC 2253
1937 PheThrLeuAspGlyArgAlaHisGlyGluLeuHisGluGlnPheLysAr 1953
2254 TTTACTTTAGATGGCCGTGGCCCATGGGGAACCTCATGAACAATTCAAAAG 2303
1953 glySThrIleLeuThrThrSerHisAlaPheProTyrIleLysThrArgV 1970
2304 GAAGACCATTTCTGACTACGCTCATGCTTTCCTTATATTAACAAGGG 2353
1970 aLAsnValThrHisLysGluGluIleIleLeuThrProIleGluValAla 1986
2354 TCAATGTCACTCATTAAGAAAGAGATCATCTTAACACCAATTGAAGTTGCT 2403
1987 lIeGluAspMetGlnLysLysThrGlnGluLeuAlaPheAlaThrHisG1 2003
2404 ATTGAGGACATGCAGAAAAAGACACAGAGATTGGCATTTGCAACACATCA 2453
2003 nasProAlaAspProLysMetLeuGlnMetValLeuGlnGlySerValG 2020
2454 GGATCCCGCAGACCCCAAAATGCTTCAGATGGTACTCCAGGATCTGTAG 2503
2020 lYThrThrValAsnGlnGlyProLeuGluValAlaGlnValPheLeuSer 2036
2504 GCACACACAGTGAATCAGGGGCTTTGGAAGTTGCCCAGGTTTTCGTCT 2553
2037 GluIleProSerAspProLysLeuPheArgHisHisAsnLysLeuArgLe 2053
2554 GAAATACCTAGTGAACCCAAAGCTCTTCAGACATCATATTAACCTGCAGCT 2603
2053 uCysPheLysAspPheThrLysArgCysGluAspAlaLeuArgLysAsnL 2070
2604 CTGCTTTAAAGATTTTACTTAAAGGTGTGAAGATGCCCTTAAGAAAAATA 2653
2070 ySSerLeuIleGlyProValGlnLysGluTyrGlnArgGluLeuGlyLys 2086
2654 AGAGCTTAATGGCGCCGTTCAAAAGAGATATCAAAAGGAATTGGGAAAA 2703
2087 leuSerSerPro 2090
2704 CTATCTTCGCCCT 2715
seq_name: /SIDSl/gcgdata/hold-geneseq/geneseqn-emb1/NA2001B.DAT:ABL17737
seq_documentation_block:
ID ABL17737 standard; DNA; 5589 BP.
XX
AC ABL17737;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 4684.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ds.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
XX
PI 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
DR WPI; 2001-656860/75.

XX	New isolated nucleic acid detection reagent for detecting 1000 or more
PT	genes from Drosophila and for elucidating cell signalling and cell-cell
PT	interactions -
XX	
PS	Claim 1; SEQ ID NO 4684; 21pp + Sequence Listing; English.
XX	
CC	The invention relates to an isolated nucleic acid detection reagent
CC	capable of detecting 1000 or more genes from Drosophila. The invention is
CC	useful in developmental biology and in elucidating cell signalling and
CC	cell-cell interactions in higher eukaryotes for the development of
CC	insecticides, therapeutics and pharmaceutical drugs. The invention
CC	discloses genomic DNA sequences (ABL01840-ABL16175) and the encoded proteins
CC	sequences (ABB01840-ABB72072).
CC	(ABB57737-ABB72072).
CC	The sequence data for this patent did not form part of the printed
CC	specification, but was obtained in electronic format directly from WIPO
CC	at ftp.wipo.int/pub/published_pct_sequences.
XX	
SEQ	Sequence 5589 BP; 1728 A; 1193 C; 1124 G; 1544 T; 0 other;

[illegible]

```
alignment_block:
US-09-737-246-2 x ABL17737 ..
```

Align seg 1/1 to: ABL17737 from: 1 to: 5589

4	ArgAlaAlaPheAlaGlnLysIleSerArgThrValAlaAlaGlnVal	20
113	CAGCGCTGTTGTGCAAAAAGTTTCCAAGCAGTCGGCTGGGATGTACG	162
20	glycGlnIleSerGlyGlnTyrSerGlySerProGlnLeuLeuLysAsnL	37
163	AAAAATGTTTCTGTTGCCAT	184
37	euaSnIleValGlyAsnIleSerHisThrThrValProLeuThrGlu	53
185CFATCAAG	193
54	AlaValAsp.....Pro	58
194	GCAATGCAGCCCATCATGTGTGTCTTCGATCTCCCGACAGAGCCCT	243
58	IAspLeuGluAspTyrLeuIleThrHisProLeuAlaValAspSerGlyP	75
244	TGATTACGAAGAATTCTATCCACGATATGAACATCATCAACCGGATC	293
75	tolLeuArgAspLeuIleGluPheProAspAspIleGluValValTyr	91
294	CCTTGAACATATATATTGACATTCTCAAGGGGATGTAACTGTAAATC	343
92	SerProArgAspCysArgThrLeuValSerAlaValProGluGlu.....	106
344	ATTTCCCGAAAAATCAGAACTGTGATCACGTTGTTCCCAATGAGAAATT	393
107	SerGluMetAspProHisValArgAspCysIleArgSerTyrThrGluA	123
394	GTTGTGATTGTCCCATCAGCTACAGAAATGTGTTAACTGCTATACGCCGC	443
123	sprTpAlaIleVal.....	127
444	CATGGAAGGTGTTGAATATGCACAGCGCACCTTTCAGTTCTTGCTAT	493
128	IleArgLysTyrHisLysLeuGlyThrGlyPheAsnProAsnThrLeuAs	144
494	ATCCGCGAGCGTATCGACCGCGGAAC..ATAAGTCCCTCTGCGGTACCA	540
144	plysGlnLysGlnArgGlnLysGlyLeuProLysGlnValPheGluSerA	161

```

541 GCAGGAATTCGAGATCGACAAGGAT ..... TTTTCGCAT 575
161 spGlAlaProAspGlyAsnSerTyrglnAspAspGln..AspAspLeu 176
576 TTGA....GAGCGTTCGCTTATAAAGCGAGAGAGCTGCATCCAGT 619
177 LysArgArgSerMet.....SerIleAspAsp.. 185
620 TCTCGACAAATCCATTGCAAGTTGGCTTCAGTAAGCTCTGTGACGACAC 669
186 ...ThrProArgGlySerTrrAlaCysSerIlePheAspLeuLysAsn 201
670 TTTGACCCCCCGCTGTTCTTGGGCTAAC...TTGATCTTGCACGTT 713
201 erLeuProAspAlaLeuLeuProAsnLeuLeuAspArgThrProAsnGlu 217
714 CGGTAAATGACCCACTAATACCAATTTGCTTGACAATGTACCTCCGAG 763
218 GluIleAspArgGlnAsnAspAspGlnArgLysSerAsnArgHisLysG 234
764 CATATTGATCAATCCATATGACGACGCCGCCGACAGATCGACAAATGCC 813
234 uLeuPheAlaLeuHisProSerProAspGluGluGluProIleGluArgL 251
814 TCTTTTCTCACTGTATCCGGAATCTGAAGCGGAGGAAACATTTGAACGCC 863
251 euSerValProAspIleProLysGluHisPheGlyGlnArgLeuVal 267
864 GCCTGCTCGCAGAGATTCCAGTTGAACATATGGTCAATCGCATTCAGGTT 913
268 LysCysLeuSerLeuLysPheGluIleGluIleGluProIlePheAlaSe 284
914 AACTGCCCTACAGCTTCGGCTTGAGCTGGAAGTGAGCCAAATTTGCATC 963
284 rLeuAlaLeuTyrrAspValLysGluLysLysLysIleSerGluAsnPhet 301
964 GATGGCTATTATGATGCTAAGGAACGGCAAAATTTCTGAATAATTCT 1013
301 yrrPheAspLeuAsnSerGluGlnMetLysGlyLeuLeuArgProHisVal 317
1014 ATTTTGACATGAACGCAGATAGTCTTAACGCATGTGTCTCAAGTCATGTA 1063
318 ProProAlaAlaIleThrThrLeuAlaArgSerAlaIlePheSerIleTh 334
1064 CAGTGCGCCGACATAGTACCAGAGCCACTCAGCCATATTCGAAATTTTC 1113
334 rTyrrProSerGlnAspValPheLeuValIleLysLeuGluLysValLeuG 351
1114 ATATCCAGTAACGATTATTCTCGTGAATTCGTCTGAAAAAGCTTTTG. 1162
351 lngInglAspIleGlyGluCysAlaGluProTyrrMetIlePheLysglu 367
1163 ..CAGGGTGACATAAACCAATCCGTCGAGCCGATATCTC..... 1198
368 AlaAspAlaThrLysAsnLysGluLysLeu..GluLysLeuLysSerG 383
1199 .....AAGAAGACAAAGATAAATAATACCGAGAGAAAGTAATAATCCAA 1239
383 nAlaAspGlnPheCysGlnArgLeuGlyLysTyrrArgMetProPheAlaT 400
1240 CGCTGTGACTACTGTGAACGACTAGGCAATAATACCGAATGCCGTTTGCAT 1289
400 rPThrAlaIleHisLeuMetAsnIleValSer..... 410
1290 GGACAGCAATCTACCTAACAAATGTTTTTAATGGCGATACTTTGAAGC 1339
411 ..SerAlaGlySerLeuGluArgAspSer..... 419
1340 AAAGACGCTGCGCCGCGGAGCGGATTCCTTTGGAATGCTTGTCGACAG 1389
420 .....ThrGluValGluIleSerThrGlyGluArgLysGlySerT 433

```


4660	4660
1614	ILELeuThrTyrAlaGluGluAspLeuGluLeuArgGluThrPheR	1630
4660	4660
1630	oAspGlnValGlnAspLeuValPheAsnLeuHisMetIleLeuSerAspT	1647
4660	4660
1647	hrValLysMetLysGlnHisGlnGluAspProGluMetLeuIleAspLeu	1663
4660	4660
1664	MetTyrArgIleAlaLysGlyTyrGlnThrSerProGluArgLeuThrTr	1680
4660	4660
1680	pleuGlnAsnMetAlaGlyLysHisSerGluArgSerAsnHisAlaGluA	1697
4660	4660
1697	IaAlaGlnCysLeuValHisSerAlaAlaLeuValAlaGluTyrLeuSer	1713
4660	4660
1714	MetLeuGluAspArgLysTyrLeuProValGlyCysValThrPheGlnAs	1730
4660	4660
1730	nIleSerSerAsnValLeuGluGluSerAlaValSerAspAspValValS	1747
4660	4660
1747	erProAspGluGluGlyIleCysSerGlyLysTyrPheThrGluSerGly	1763
4660	4660
1764	LeuValGlyLeuLeuGluGlnAlaAlaAspPheSerMetAlaGlyMe	1780
4660	4660
1780	tTyrGluAlaValAsnGluValTyrLysValLeuIleProIleHisGluA	1797
4660	4660
1797	IaAsnArgAspAlaLysLysLeuSerThrIleHisGlyLysLeuGlnGlu	1813
4660	4660
1814	AlaPheSerLysIleValHisGlnSerThrGlyTyrGluArgMetPheG1	1830
4660	4660
1830	yThrTyrPheArgValGlyPheTyrGlyThrLysPheGlyAspLeuAspG	1847
4660	4660
1847	LuGlnGluPheValTyrLysGluProAlaIleThrLysLeuAlaGluIle	1863
4660	4660
1864	SeuHisArgLeuGluGlyPheTyrGlyLysArgPheGlyGluAspValVa	1880
4661TTTAACTGACGATTCGACCGAGCTCTGT	4692
1880	IGluValIleLysAspSerAsnProValAspLysCysLysLeuAspProA	1897
4693	GCATATCATTTAAAGATTCCAATACCGTTGATATTAAATAGCTTGATCCCG	4742
1897	snLysAlaTyrIleGlnIleThrTyrValGluProTyrPheAspThrTyr	1913
4743	ATTAAGGCTTACATTCAAATTACTTATGTTGAACCTACTTTGAACATAT	4792

```

1914 GluMetLysAspArgIleThrTyrPheAspLysAsnTyrAsnLeuArgAr 1930
|||||::: ||| |||||||:::|||||:::|||||:::|||||
4793 GAAATGCGTCATCGTAGACATACTTTGAGCGGAATTTCAATATAAACG 4842
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
1930 gPheMetTyrCysThrProPheThrLeuAspGlyArgAlaHisGlyIul 1947
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
4843 TTTTATATATGCGACGCGCTTTTACTAAAAATGGAAGGCACATGGCGAAC 4892
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
1947 euhHisGluGlnPheLysArgLysThrIleLeuThrThrSerHisAlaPhe 1963
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
4893 TAAATGAGCAGTGTAAACGGAAACAGATTTTGACACAGCGCAATCACATT 4942
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
1964 ProTyrIleLysThrArgValAsnValThrHisLysGluGluIleIle 1980
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
4943 CCATACGTAAACATCCGATTTATGTAATTAGCCGACAGCAATAGTCTCT 4992
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
1980 uThrProIleGluValAlaIleGluAspMetGlnLysLysThrGlnGlu 1997
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
4993 GGAACCGGATTGAAGTGGCAATTGAAGATATTCAAAAAAACATTGGAAT 5042
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
1997 euAlaPheAlaThrHisGlnAspProAlaAspProLysMetLeuGlnMet 2013
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
5043 TGGCAGCTGCCAGCATCAAGACCGACGACCCCAAAATATTGCAAAATG 5092
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
2014 ValLeuGlnGlySerValGlyThrThrValAsnGlnGlyProLeuGluVa 2030
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
5093 TACTGTCAGGATGTATTGGAACCACTGTTAACCAAGACCGATGGAAT 5142
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
2030 lAlaGlnValPheLeuSerGluIleProSerAspProLysLeu..Phea 2046
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
5143 GCGGAGTGTGTCTTCCTTCCCAATTATCCGACGCAACAACGTACCACAA 5192
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
2046 rghHisHisAsnLysLeuArgLysCysPheLysAspPheThrLysArgCys 2062
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
5193 AACACCAAAACAACATTCGGTTGTGCTTTCGCGAGTTTCAAAAGCGTTGT 5242
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
2063 GluAspAlaLeuArgLysAsnLysSerLeuIleGlyProValGlnLysGlu 2079
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
5243 GCTGATGCTTTGAAAAAGAAATCGCAATCTAATACTTTCAGATCAAAAAAGA 5292
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
2079 uTyrGlnArgGluLeu 2084
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
5293 TTACCAACGAGAACTG 5308
|||||:::|||||:::|||||:::|||||:::|||||:::|||||

seq_name: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA2001A.DAT:AAS07382
seq_documentation_block:
ID AAS07382 standard; cDNA; 4027 BP.
XX
AC AAS07382;
XX
DT 26-SEP-2001 (first entry)
XX
DE Human DNA associated with CLASP-5 #2.
XX
KW Human; CLASP-5; cadherin-like asymmetry protein; immune gateway;
KW immunogen; antibody; autoimmune disease; rheumatoid arthritis;
KW multiple sclerosis; leukaemia; insulin dependent diabetes mellitis;
KW acquired immunodeficiency syndrome; AIDS; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..2970
FT /*tag= a
FT /product= "Protein as displayed in AAU04026"
FT /partial
FT /note= "No start codon"
XX
PN WO200142296-A2.
XX
PD 14-JUN-2001.

```

```

seq_name: /SIDSl/gcgdata/hold-geneseq/geneseqn-emb1/NA2001A.DAT:AAS07382
seq_documentation_block:
ID   AAS07382 standard; cDNA; 4027 BP.
XX
AC   AAS07382;
XX
DT   26-SEP-2001 (first entry)
XX
DE   Human DNA associated with CLASP-5 #2.
XX
KW   Human; CLASP-5; cadherin-like asymmetry protein; immune gateway;
KW   immunogen; antibody; autoimmune disease; rheumatoid arthritis;
KW   multiple sclerosis; leukaemia; insulin dependent diabetes mellitis;
KW   acquired immunodeficiency syndrome; AIDS; ss.
XX
OS   Homo sapiens.
XX
FH   Key
FT   CDS
FT   1.2970
FT   /*tag= a
FT   /product= "Protein as displayed in AAU04026"
FT   /partial
FT   /note= "No start codon"
XX
PN   WO200142296-A2.
XX
PD   14-JUN-2001.

```

```

XX PF 13-DEC-2000; 2000OWO-US34163.
XX PR 13-DEC-1999; 99US-0170453.
PR 14-JAN-2000; 2000US-0176195.
PR 14-FEB-2000; 2000US-0182296.
PR 11-APR-2000; 2000US-0196267.
PR 11-APR-2000; 2000US-0196460.
PR 11-APR-2000; 2000US-0196527.
PR 11-APR-2000; 2000US-0196528.
PR 11-APR-2000; 2000US-0547276.
PR 13-OCT-2000; 2000US-0240503.
PR 13-OCT-2000; 2000US-0240508.
PR 13-OCT-2000; 2000US-0240543.
PR 13-OCT-2000; 2000US-0240539.
XX PA (ARBO-) ARBOR VITA CORP.
XX PI Lu P, Garman JD, Candia AF;
XX DR WPI; 2001-367865/38.
DR P-PsDB; AAU04026.
XX PT CLASP-5 polynucleotides, proteins and antibodies are used to prevent or
PT treat a CLASP-5 mediated disease, such as an autoimmune disease e.g.
PT Rheumatoid arthritis -
XX PS Disclosure; Fig 9C; 188pp; English.
XX CC The sequence encodes a protein presented as AAU04026, included in the
CC specification which relates to a cadherin-like asymmetry protein,
CC CLASP-5, which is a transmembrane protein of the immune system involved
CC in the formation of the immune gateway. CLASP-5, polynucleotides encoding
CC it and an anti-CLASP-5 antibody are used to prevent or treat a CLASP-5
CC mediated disease, such as an autoimmune disease caused or exacerbated by
CC increased activity of TH1 cells. These diseases may include Addison's
CC disease, rheumatoid arthritis, dermatitis, Multiple Sclerosis, Systemic
CC Lupus Erythematosus and Autoimmune Thyroiditis, inflammatory conditions
CC (e.g Ischaemia-reperfusion) and responses, Leukaemia, acquired
CC immunodeficiency syndrome, AIDS, syphilis, typhoid, Burkitt's lymphoma,
CC Grave's disease and insulin dependent diabetes mellitus. CLASP-5 is used
CC for inhibiting an immune response in a cell (T cell or B cell) by
CC interfering with the expression of a CLASP-5 gene in the cell, the
CC ability of a CLASP-5 protein to bind to another cell or the ability of a
CC CLASP-5 protein to bind to another protein. The antibody binding CLASP-5
CC is used to inhibit an immune response is a subject. The polynucleotide is
CC used to detect CLASP-5 expression in cells and for diagnosis of
CC diseases and disorders associated with aberrant expression of CLASP-5.
CC Note: The present sequence is included in the specification but is
CC not mentioned anywhere in the specification.
XX SQ Sequence 4027 BP; 1162 A; 898 C; 957 G; 1010 T; 0 other;

alignment_scores:
      Quality: 3025.00          Length: 949
      Ratio:   3.716           Gaps:    9
Percent Similarity: 85.774     Percent Identity: 62.171

alignment_block:
US-09-737-246-2 x AAS07382 ..

Align seg 1/1 to: AAS07382 from: 1 to:.4027

1142 ThrProProAlaserProSerProSerValserSerAlaThrSergInse 1158
||| |||::||||| |||||::||| |||||::
19 ACTGCTCCAACATCTCTGTCTTCATATCT.....TCCAGAA 59

1158 rSergLPheserThrAsnValGlnaspGlnLysIleAlaAsnMetPheg 1175
::::: ::::: ||||| ||||| ||||| |||||
60 CTCACAGCTCCTGCTCCAGCTTCCAGACCAGAAGATCGCCAGCATGTTCG 109

1175 luleUser.ValPrOpheargGlnGlnHisTyrrLeuAla.GlyLeuVal 1190

```

```

110 ATCGAAGCTTCCAGAGTACCGCGCCAGACGACACTTCCTCACCGGGGCTCTC 159
    : : : : | | | | | | | : : | | : : | | | | : :
1191 LeuThrGluLeuAlaValIleLeuAspProAspAlaGluGlyLeuPheG1 1207
    : : | | | | | | | : : | | | | : : : : | | | | : :
160 TTCACAGAACTGGCTGCTGCCCTGGATGCCGAAGGGGAAGGAATCAGCCA 209
    : : : : : : : : : : : : : : : : : : : : : : : :
1207 yLeuHisLysLysValIleAsnMetValHisAsnLeuLeuSerSerHis 1224
    : : : : : : : : : : : : : : : : : : | | | | | | | |
210 AGTACAAAGGAAGAGCTGTCAAGTGCATTCACAGCCCTGCTAAGTTCACG 259
    : : : : : : : : : : : : : : : : : : : : : : : :
1224 spSerAspProArgTyrSerAspProGlnIleLysAlaArgValAlaMet 1240
    | | | | | | | | | | | | | | : : : : : : : : : :
260 ACCTCGACCCACGCTGTGTCAACCCAGAGGTGAAGGTCAAAATCGCGCC 309
    : : : : : : : : : : : : : : : : : : : : : : : :
1241 LeuTyrLeuProLeuIleGlyIleIleMetGluThrValProGlnLeuTyr 1257
    | | | | | | | | | : : | | | | : : : : : : : : : :
310 CTTTACCTAACCTTTAGTTGGCATCATTTTGGATGCTTTGCCACAGCTCTG 359
    : : : : : : : : : : : : : : : : : : : : : : : :
1257 rAspPheThrGluThrHisAsnGlnArgGlyArgProIleCysIleAlaT 1274
    | | | | | | | : : : : : : | | : : : : : : : :
360 TGACTTTACAGTTGCAGATACTCGCAGATACCCG.....ACCAATG 400
    : : : : : : : : : : : : : : : : : : : : : : : :
1274 hrAspAspTyrGluSerGluSerGlySerMetIleSerGlnThrValAla 1290
    : : : : : : : : : : : : : : : : : : : : : : : :
401 GCTCGGATGAAGAACAGAAGAGAGCCGGTCCCATTTACCCAGAATGTGCT 450
    : : : : : : : : : : : : : : : : : : : : : : : :
1291 MetaIleAlaGlyThrSerValProGlnLeuThrArgProGlySerPh 1307
    : : | | | | | | | : : : : : : | | : : : : : :
451 CTGGCCATAGCAGGAATATTTCATTGAAACA...AGTGAAT 494
    : : : : : : : : : : : : : : : : : : : : : : : :
1307 eleuLeuThrSerThrSerGlyArgGlnHisThrThrPheSerAlaGluS 1324
    : : : : : : : : | | : : | | : : : : : : : : : :
495 AGTGTGTCTTCCTTGCCCTATAGCAGTACAAACATGCTGAACGCGGACA 544
    : : : : : : : : : : : : : : : : : : : : : : : :
1324 erSerArgSerLeuLeuIleCysLeuLeuThrValLeuLysAsnAlaAsp 1340
    : : : : | | : : : : : : : : : : : : : : : : | | | |
545 CTACTCGCAACCTCATGATCTGCTCTCTGGATCATGAATAATGCTGAT 594
    : : : : : : : : : : : : : : : : : : : : : : : :
1341 GluThrValLeuGlnLysTrpPheThrAspLeuSerValLeuGlnLeuAs 1357
    : : : : : : : : : : | | | | : : : : | | | |
595 CAGAGCCTCATTAGGAAGTGGATGTGTCACCTGCCATCAACGCACGCTCAA 644
    : : : : : : : : : : : : : : : : : : : : : : : :
1357 nArgLeuLeuAspLeuLeuTyrIleCysValSerCysPheGluTyrLysG 1374
    | | | : : | | | | | | : : : : | | | | | | | |
645 CAGGATTTTAGATCTACTTTTCATCTGTGTATGTTTGGATATTAAGG 694
    : : : : : : : : : : : : : : : : : : : : : : : :
1374 LyLysLysValPheGluArgMetAsnSerLeuThrPheLysLysSerLys 1390
    | | | : : : : : : : : : : : : : : : : : : : : :
695 GAAAAACAGAGTTCTGACAAAGTCACTAACCCAAAGTCTGCAGAAAGTCAA 744
    : : : : : : : : : : : : : : : : : : : : : : : :
1391 AspMetArgAlaLysLeuGlnGluAlaIleLeuGlySerIleGlyAlaAr 1407
    : : : : : : : : : : : : : : : : : : : : : : : :
745 GATGTCAAGGCCCGGCTGGAAGAGCTTGTCTGCGTGGGAAGGGCCAG 794
    : : : : : : : : : : : : : : : : : : : : : : : :
1407 gGlnGluMetValArgArgSerArgGlyGlnLeuGluArgSerProSerG 1424
    | | | | | | : : | | | | : : : : : : : : : :
795 AGGGGAGATGATGCGCGC.....CGGGCTCCAGGGA 826
    : : : : : : : : : : : : : : : : : : : : : : : :
1424 LySerAlaPhe...GlySerGlnGluAsnLeuArgTrpArgLysAspMet 1439
    : : : : | | | | : : | | | | | | : : : : : :
827 ACGACCGCATTTCCAGGCCCTAAATGAATAATTGAGATGGAAAGAAAGCAG 876
    : : : : : : : : : : : : : : : : : : : : : : : :
1440 ThrHisTrpArgGlnAsnThrGluLysLeuAspLysSerArgAlaGluI1 1456
    | | | | | | | | : : | | | | | | : : : : : :
877 ACACATTGGCGGCAAGCTAATGAGAACCTAGATAAAAACAAAGGCCGAGTT 926
    : : : : : : : : : : : : : : : : : : : : : : : :
1456 eGluHisGluAlaLeuIleAspGlyAsnLeuAlaThrGluAlaAsnLeu 1473
    : : : : : : | | | | : : | | | | | | : : : :
927 AGATCAAGAAAGCCTTGATCATGTGCATCTGGCTACAGAAGCACATTTAA 976
    : : : : : : : : : : : : : : : : : : : : : : : :
1473 leIleLeuAspThrLeuGluIleValValGlnThrValSerValThrGlu 1489
    | | | | | | | | | | : : : : : : : : : : : :

```

[illegible]

1789	LysValLeuIleProIleHisGluAlaAsnArgAspAlaLysLysLeuSe	1805
1927	AACTGGTCATCCCATCTAGAAAGCGCATGAGAAATCCGGAAGCTGAC	1976
1805	rThrIleHisGlyLysLeuGlnGluAlaPheSerLysIleValHisGlnS	1822
1977	ACTTCACCTCACAGCAAGCTGCAGAGAGCCCTTCGACAGCATCGTTAAACAAGG	2026
1822	erThrGlyTrpGluArgMetPheGlyThrTyrPheArgValGlyPheTyr	1838
2027	AT.....CATAGAGAATGTTTGAACCTACTTCCGACTTGTTCTT	2070
1839	GlyThrLysPheGlyAspLeuAspGluGlnGluPheValTyrLysGluPr	1855
2071	GGATCCAAATTTGGGATTTGGATGAACAGAGATTGTCTACAAAGAGCC	2120
1855	oAlaIleThrLysLeuAlaGluIleSerHisArgLeuGluGlyPheTyrG	1872
2121	TGCATTTACCAAGCTTCCTGAGATCTCACATAGACTAGAGCATTTTATG	2170
1872	lyGluArgPheGlyGluAspValAlaGluValIleLysAspSerAsnPro	1888
2171	GTCAATGTTTGGTGCAGAAATTTGTGGAAGTGATTAAGACTCCACTCCT	2220
1889	ValAspLysCysLysLeuAspProAsnLysAlaTyrIleGlnIleThrTy	1905
2221	GTGCACAAACCAAGTTGGATTCCTAACAAAGCCCTACATCAGATCACTTT	2270
1905	rValGluProTyrPheAspThrTyrGluMetLysAspArgIleThrTyrP	1922
2271	TGTGAGACCCTACTTTGATGAGATGAGATGAAGAGACAGGTCACATACT	2320
1922	heAspLysAsnTyrAsnLeuArgArgPheMetTyrCysThrProPheThr	1938
2321	TTGAGAAGAATTTCAACCTCCGAGGTTTCATGTACACACCACCCGTTCCAC	2370
1939	LeuAspGlyArgAlaHisGlyGluLeuHisGluGlnPheLysArgLysTh	1955
2371	CTGAGAGGGGGCGCCCTCGGGGAGAGCTGCATGAGCAGTACAGAAGAACAC	2420
1955	rIleLeuThrThrSerHisAlaPheProTyrIleLysThrArgValAsnV	1972
2421	AGTCTGACCACCTATGCACGCCCTCCCTACATCAAGACAGGATCAGCG	2470
1972	AlpThrHisLysGluGluIleIleLeuThrProIleGluValAlaIleGlu	1988
2471	TCAATCCAGAAGAGAGAGTTGTTTGGACACCGATTGAAGTTGCCATTGAA	2520
1989	AspMetGlnLysLysThrGlnGluLeuAlaPheAlaThrHisGlnAspPr	2005
2521	GACATGAAGAAGAAGACCCTGCAGTTAGCAGTTGCCATTAAACCAAGAGCC	2570
2005	oAlaAspProLysMetLeuGlnMetValLeuGlnGlySerValGlyThrT	2022
2571	GCCTGATGCAAAAGATGCTTCAGATGGTGCTGCAAGGCTCTGTGGAGCTPA	2620
2022	hrValAsnGlnGlyProLeuGluValAlaGlnValPheLeuSerGluIle	2038
2621	CTGTAAATCAGGAGCACCTGGAGTAGACCCCAAGTGTTTTGCGCTGAATT	2670
2039	ProSerAspProLysLeuPheArgHisHisAsnLysLeuArgLeuGlyCysPh	2055
2671	CCTGCTGATCCAAACTCTATTCGACATCACACACAGTTGAGCTTATGCTT	2720
2055	eLysAspPheThrLysArgCysGluAspAlaLeuArgLysAsnLysSerL	2072
2721	TAAAGAAATTCATCATGAGATGTGTGAAGCTGTAGAGAAAAACAAGCGTC	2770
2072	euIleGlyProValGlnLysGluTyrGlnArgGluLeuGlyLys	2086
2771	TCATTCACGGCAGACACAGAGGGAATATTCAGCAGGAAGACTCAAAAAG	2814

2072 eurlleelyProvaIglInlysgIuTyrgInargIuIeugIyLys 2086
 ||||| |||:::|||||||:::||||| |||
 2771 TCATCAGGGCAGACCAGAGGGATATCAGCAGGAACCTCAAAAG 2814

```
seq_name: /SIDS1/gcgcdata/hold-geneseq/geneseqn-emb1/NA2001A.DAT:AAH43850
seq_documentation_block:
ID   AAH43850 standard; cDNA; 2148 BP.
XX
AC   AAH43850;
XX
DT   04-SEP-2001 (first entry)
XX
DE   Preliminary human CLASP-7 encoding cDNA sequence.
XX
KW   Human; CLASP-7; cadherin-like asymmetry protein; immune response;
KW   neuroprotective; antidiabetic; immunosuppressive; antirheumatic;
KW   antiarthritic; hypotensive; anti-HIV; cytostatic; immunostimulant;
KW   antianaemic; antinflammatory; ophthalmological; nephrotrophic;
KW   antithyroid; antiasthmatic; antiallergic; antibacterial; gene therapy;
KW   chromosome 19q13.2; autoimmune disease; multiple sclerosis; toxemia;
KW   juvenile diabetes; rheumatoid arthritis; pruritic urticarial papule;
KW   hypertension; Rh incompatibility; ss.
XX
OS   Homo sapiens.
XX
FH   Key Location/Qualifiers
FT   CDS 1..1932
FT             /tag= a
FT             /partial
FT             /product= "preliminary CLASP-7 protein"
FT             /note= "no start codon given"
XX
PN   WO200142295-A2.
XX
PD   14-JUN-2001.
XX
PF   13-DEC-2000; 2000WO-US34152.
XX
PR   13-DEC-1999; 99US-0170453.
PR   14-JAN-2000; 2000US-0176195.
PR   14-FEB-2000; 2000US-0182296.
PR   11-APR-2000; 2000US-0196267.
PR   11-APR-2000; 2000US-0196460.
PR   11-APR-2000; 2000US-0196527.
PR   11-APR-2000; 2000US-0196528.
PR   11-APR-2000; 2000US-0547276.
PR   13-OCT-2000; 2000US-0240503.
PR   13-OCT-2000; 2000US-0240508.
PR   13-OCT-2000; 2000US-0240539.
PR   13-OCT-2000; 2000US-0240543.
XX
PA   (ARBO-) ARBOR VITA CORP.
XX
PI   Lu P, Garman JD, Candia AF;
XX
DR   WPI; 2001-381641/40.
DR   P-PSDB; AAB99540.
XX
XX
PT   Novel cadherin-like asymmetry protein-7 and polynucleotides encoding
PT   the polypeptide, useful for treating autoimmune disease,
PT   hypersensitivity, preventing transplant rejection by modulating immune
PT   response -
XX
PS   Example 3; Fig 1; 151pp; English.
XX
CC   The present invention describes a human cadherin-like asymmetry protein
CC   (CLASP), designated CLASP-7. The CLASP-7 protein (I) and its encoding
CC   nucleotide sequence (II) have activities including: neuroprotective;
CC   antidiabetic; immunosuppressive; antirheumatic; antiarthritic; anti-HIV;
CC   hypotensive; cytostatic; immunostimulant; antianaemic; antinflammatory;
CC   ophthalmological; nephrotrophic; antithyroid; antiasthmatic;
CC   antiallergic; and antibacterial. (II) and CLASP-7 antibodies (III) are
CC   useful for detecting the CLASP-7 polypeptide. (II) is useful for
CC   producing (I) by recombinant methods. (I) or its fragment are useful for
CC   inhibiting an immune response in a cell such as T cell or B cell. A
CC   pharmaceutical composition (C), comprising (I) or (II), can be useful
```

```
CC   for treating CLASP-7-mediated disease such as an autoimmune disease
CC   caused or exacerbated by increased activity of T helper cells. Autoimmune
CC   diseases which can be treated using (C) include multiple sclerosis,
CC   juvenile diabetes and rheumatoid arthritis. (I) is useful for treating
CC   toxemia or pregnancy induced hypertension, pruritic urticarial papules
CC   and Rh incompatibility. (I) is also useful as a diagnostic reagent for
CC   immune and other disorders, since diseases characterised by
CC   overproduction or depletion of lymphocytes in blood or other organs may
CC   be detected by monitoring the level of (I) or its mRNA. CLASP-7 has been
CC   mapped to the chromosomal location 19q13.2. The present sequence
CC   represents the preliminary human CLASP-7 encoding cDNA sequence which is
CC   given in the present invention.
XX
SQ   Sequence 2148 BP; 468 A; 649 C; 625 G; 406 T; 0 other;
XX
alignment_scores:
Quality: 2552.50 Length: 599
Ratio: 4.534 Gaps: 1
Percent Similarity: 93.990 Percent Identity: 80.467
alignment_block:
US-09-737-246-2 x AAH43850 ..
Align seg 1/1 to: AAH43850 from: 1 to: 2148
1487 ValThrGluSerLysGluSerIleLeuGlyValLeuLysValLeuE 1503
      ::::::::::::::::::::::::::::
1 CTTTCAGAGAAGCCCCGGGAGAGCGCTCTGGGGGATTGTGCTGAAGTTGTGCT 50
1503 uHisSerMetalAcysasnGlnSerAlaValTyrLeuGlnHisCysPheA 1520
      ::::::::::::::::::::
51 GTACAGCCTGGGCAGTGCCTCCAGAGTGCCTCTTCTTGACGATGGCCTGG 100
1520 lAThrGlnArgAlaLeuValSerLysPheProGluLeuPheGluGlu 1536
      ::::::::::::::::::::::::::::::::::::
101 CCACCCAGAGGGCCCTGTGTCCAAAGTCCCGGAGCTGCTGTCGAGGAG 150
1537 GluThrGluGlnCysAlaAspLeuCysLeuArgLeuLeuArgHisCysSe 1553
      ::::::::::::::::::::::::::::::::::::
151 GACACGGAGCTGTGTGCGGACCTGTGCTGAGGCTCCTACGACACATGTG 200
1553 rSerSerIleGlyThrIleArgSerHisProSerAlaSerLeuTyrLeuL 1570
      ::::::::::::::::::::::::::::
201 CAGCCGCATCAGCACCATCCCGCACGACGCCAGCGCCTCGTGTACTGCG 250
1570 euMetArgGlnAsnPhcGluIleGlyAsnAsnPheAlaArgValLysMet 1586
      ::::::::::::::::::::::::::::::::::::
251 TCATGCGAGAGAACTTGAGATCGGCCACACACTTTGCCCGTGAAGATG 300
1587 GlnValProMetSerLeuSerSerLeuValGlyThrSerGlnAsnPheAs 1603
      ::::::::::::::::::::::::::::::::::::
301 CAGGTCAACCATGTCTCTCTGCTCCCTGCTGGGGAGCAGCAGAACTTCAG 350
1603 nGluGluPheLeuArgArgSerLeuLysThrIleLeuThrTyrAlaGluG 1620
      ::::::::::::::::::::::::::::::::::::
351 TGAAGAGCACCCTGCGAGCTTCACCTCAAAACCATCCTCACCTATGCTGAG 400
1620 lAspPLeuGluLeuArgGluThrThrPheProAspGlnValGlnAspLeu 1636
      ::::::::::::::::::::::::::::
401 AGGACATGGGGCTGCGGGAGACAGCACCTTCGAGAGCAGAGTCCAGACTTG 450
1637 ValPheAsnLeuHisMetIleLeuSerAspThrValLysMetLysGluHt 1653
      ::::::::::::::::::::::::::::::::::::
451 ATGTTCAACCTGCACATGATCTCTGACGAGACACGCGTGAAGATGAAGACA 500
1653 sGlnGluAspProGluMetLeuIleAspLeuMetTyrArgIleAlaLysG 1670
      ::::::::::::::::::::::::::::::::::::
501 CCAGGAGGACCTGAGATGCTCATGACCTCATGTACAGAAATTGCCGGG 550
1670 lYTYrGlnThrSerProGlu..ArgLeuThrTrpLeuGlnAsnMetaLa 1685
      ::::::::::::::::::::::::::::
551 GCTACACGAGGCTCACCGGACCTTTCGGCTGACTGTGTCAGAAACATGCCC 600
```


1686	GlyLysHisSerGluArgSerAsnHisAlaGluAlaGlnCysLeuVal	1702
601	GGGAAGCACGCCGAGCTGGGCACCAACGCCCGAGGCCGCCCACTGCATGGT	650
1702	HisSerAlaAlaLeuValAlaGluTyrLeuSerMetLeuGlnAspArgL	1719
651	GCACGCCGCCCTCGTGCTGAGTACCCTCGCCCTGCTCGAGGACCACC	700
1719	ySTyrrLeuProValGlyCysValThrPheGlnAsnIleSerSerAsnVal	1735
701	GCCACCTGCCCGTGGGTGGTTCTTCAGAACATCTCATCCAATGTG	750
1736	IeuGlIngluSerAlaValSeraspSpValValSerProaspGluGlu	1752
751	CTAGAGGAGTCCGCCATCTCCAGACACATCTGTCCGCCGACGAGGAGG	800
1752	yIleCyssSerGlyLysTyrPheThrGluSerGlyLeuValGlyLeuLeug	1769
801	CTTCTGCTCCGGGAGACACTTCACGTAGCTGGGGCTGTAGGGTTGCTGG	850
1769	IuGlnAlaAlaAlaSerPheSerMetAlaGlyMetTyrGluAlaValasn	1785
851	AACAGGCAGCCGGCTACTTCAACCATGGCGGGCTCTACGAGCGGTGAAT	900
1786	GluValTyrLysValLeuIleProIleHisGluAlaAsnArgaspAlaly	1802
901	GAGGTCTACAAGAACCTCATCCCCATCCTGGAAGCCCACCGTAGCTACAA	950
1802	sLySLeuSerThrIleHisGlyLysLeuGlnGluAlaPheSerLysIlev	1819
951	GAAGCTGGCCGGGTGCACGGCAAACGTGACGAGGCGCTTCAACCAAGATCA	1000
1819	aHISGLNserThrGlyTrpGluArgMetPheGlyThrTyrPheArgVal	1835
1001	TGCACACAGATTCCGGCTGGGAGCGCGTGTTCGGAGCATTTCCGCGTG	1050
1836	GlyPheTyrGlyThrLysPheGlyAspLeuAspGluGlnGluPheValTy	1852
1051	GGCTTCTACGGCGCCCACTTCGGTGAACCTGATGAGCAGAGATTGTGTA	1100
1852	rLySGluProAlaIleThrLysLeuAlaGluIleSerHisArgLeuGlu	1869
1101	CAAGGAGCCATCGATCACGACGCTGGCAGAGATCTCACACCGCGCTGAGG	1150
1869	LypHeTyrGlyGluArgPheGlyGluAspValValGluValIleLysasp	1885
1151	AGTTCACACGAGAGATTGGCGACGACGAGCTGTTGAGATTATCAAAAGAC	1200
1886	SeraspProValasplyscysLysLeuaspproAsnLysalaTyrIlegl	1902
1201	TCTTACCCTGTGACAAAGTCCAAAGCTTGACTCACAAAAGGCTACATCCA	1250
1902	nIleThrTyrValGluProtyrPheasptHrtyrGluMetLysaspArgI	1919
1251	GATCACGTAATGTGAACCGTAATTGTATACCTACGAGCTCAAGACCGGG	1300
1919	IethrTyrPheaspLysasnTyrAsnLeuArgArghemetyrcystThr	1935
1301	TGACCTACTTTGACCGCACTATGGGCTTCGCACATTCCTGTCTGCACG	1350
1936	ProPhetHrLeuAspGlyArgAlaHisglyGluLeuHisGluGlnPhely	1952
1351	CCGTTCAACGCCGATGGCGCGCACACAGGGGAGCTGCCCGAGCAACACAA	1400
1952	sArgLysThrIleLeuthrThrSerHisAlaPheProTyrIleLysThra	1969
1401	GCGTAAGACGCTGCTACAGCACCGAACCACGCCCTTCCCTACATCAAGACTC	1450
1969	rgValasnaValThrHisLysGluGluIleIleLeuthrProileGluVal	1985
1451	GCAATCCGTGTGCCACCGGAGAGAGACGGTGTCTGACGCCAAGTGGAGGTG	1500

1986 AlaIleGluAspMetClnLysLysThrGlnGluLeuAlaPheAlaThrHi 2002
|||||
1501 GCCATCGAGGACATGCAGAAGAAGACACGGGAGCTGGCCCTTTGCCACCGA 1550
2002 sGlnAspProAlaAspProLysMetLeuGlnMetValLeuGlnGlyServ 2019
:|||||
1551 GCAGGACCCACACAGATGCTTAAGATGTACAGATGGTGCTTCAGGGCTCTG 1600
2019 aIGlyThrThrValAsnGlnGlyProLeuGluValAlaGlnValPheLeu 2035
|||||
1601 TAGGGCCACCGTGAACCCAGGGTCCCCCTGGAGGTGGCCACAGGTGTTTTTA 1650
2036 SerGluIleProSerAspProLysLeuPheArgHisHisAsnLysLeuAr 2052
:::|||||
1651 GCAGAGATCCCCGAGAGACCCCAAGCTCTTCGGGATCACAACAATTGCG 1700
2052 gLeuCysPheLysAspPheThrLysArgCysGluAspAlaLeuArgLysA 2069
|||||
1701 GCTCTGCTTCAAGACTTCTGCAAGAATAATGTAGAGATGCGCTGCGGAATA 1750
2069 snLysSerLeuIleGlyProValGlnLysGluTyrGlnArgGluLeu 2084
|||||
1751 ATAAGCCCTGATTTGGGCGCGGACACAGAAGAGATACACCGTGAGCTG 1797
seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA2001A.DAT:AAS08334

seq_documentation_block:	
ID	AAS08334 standard; cDNA; 6454 BP.
XX	
AC	AAS08334;
XX	
DT	23-OCT-2001 (first entry)
XX	
DE	Human cDNA encoding CLASP-4.
XX	
KW	Human; CLASP-4; cadherin-like asymmetry protein-4; autoimmune disease;
KW	ataxia telangiectasia; Human immunodeficiency virus infection;
KW	inflammatory disease; rheumatoid arthritis; multiple sclerosis;
KW	diabetes mellitus; immune disorder; Guillain-Barre syndrome;
KW	severe combined immunodeficiency; allergic reaction; asthma;
KW	immunogen; antibody; ss.
XX	
OS	Homo sapiens.
XX	
FH	Key
FT	CDS
FT	Location/Qualifiers
FT	95..6121
FT	/*tag= a
FT	/product= "CLASP-4"
FT	replace (100,T)
FT	/*tag= b
FT	/note= "Missense mutation"
FT	replace (1095,G)
FT	/*tag= c
FT	/note= "Missense mutation"
FT	replace (1351,G)
FT	/*tag= d
FT	/note= "Missense mutation"
FT	replace (1628,G)
FT	/*tag= e
FT	/note= "Changes Thr to Ala"
FT	replace (2019,G)
FT	/*tag= f
FT	/note= "Changes Asn to Ser"
FT	replace (2421,A)
FT	/*tag= g
FT	/note= "Changes Gly to Asp"
FT	replace (2431,G)
FT	/*tag= h
FT	/note= "Missense mutation"
FT	replace (2439,C)
FT	/*tag= i
FT	/note= "Changes Gly to Glu"
FT	replace (2480,T)
FT	allele

764 sNAsnLeuGlUAsnGlULeuLysSerSerIleSerAlaLeuAsnSerSer 780
 :::||||| ::: :::|||||::: :::
2435 CCAGGGGAGCTCATTAATATTAAAGTGTTCATGCCATG 2476
781 GlnLeuGlUProValValArgPheLeuHisLeuLeuAspLysLeuI 797
 :::||||| :::||||| :::||||| :::|||||:::
2477 GAGATCCAAGTCATGATACAGTTCTTACCTGTAATTCCTTAAGCACTCTT 2526
797 eLeuLeuValIleArg.....ProProValIleAlaG 808
 :::||||| ::: :::||||| ::: :::
2527 CCGAGTTCACAAATATGACCCATGAGATGACGTTCCCTATCACTGCA 2576
808 LylGlnIleValAsnLeuGlLysIleAlaSerPheGluAlaMetAlaSerIle 824
 :::||||| ::: :::||||| ::: :::
2577 CCATGGT.....CTCTTACATATT 2596
825 IleAsnArgLeuHisLysAsnLeuGlUcLysAsnHisAspGlnHisGlyAr 841
 :::||||| ::: :::||||| ::: :::
2597 GTATCAAGTGCAT.....GAAGAAGGCTT 2622
841 gAsnSerLeuLeuAlaSerTyrIleHisTyrValPheArgLeuProAsnT 858
 :::||||| :::||||| :::||||| :::||||| :::
2623 GGATAGTTATCTAAGATCATTCATTAAGTATAGCTTCCGA..... 2662
858 hTyrProAsnSerSerSerProGlyProGlyGlyLeuGlyGlySerVal 874
2662 2662
875 HisTyrAlaThrMetAlaArgSerAlaValArgProAlaSerLeuAsnIle 891
2662 2662
891 uAsnArgSerArgSerLeuSerAsnSerAsnProAspIleSerGlyThrP 908
2662 2662
908 roThrSerProAspAspGluValArgSerIleIleGlySerLysGlyLeu 924
 :::||||| ::: :::||||| ::: :::
2663CCTGAA..... 2668
925 AspArgSerAsnSerTyrValAsnThrGlyGlyProLysAlaAlaProTr 941
2668 2668
941 pGlySerAsnProSerProSerAlaGluSerThrGlnAlaMetAspArgS 958
 :::||||| ::: :::||||| ::: :::
2669AAACCGAGTGT..... 2680
958 erCysAsnArgMetSerSerHisThrGluThrSerSerPheLeuGlnThr 974
2680 2680
975 LeuThrGlyArgLeuProThrLysLysLeuPheHisGluGluLeuAlaIle 991
 :::||||| :::||||| :::||||| :::||||| :::
2681CCTCAGGCCAGCTGATACATGAACCCCTGGCTAC 2715
991 uGlnTyrValValCysSerGlySerValArgGluSerAla..... 1004
 :::||||| :::||||| :::||||| ::: :::
2716 TAGCATGATAGCAATA.....TTGAACAAGTCTGCAGATTTTATTAT 2756
1005LeuGlnGlnAlaTyrPhePhePheGluLeuMetVal 1016
 :::||||| :::||||| :::||||| :::||||| :::
2757 CAATTAACAATTCCTAAAGTACTCATGCTTTTCTTTGAATTAATTGCA 2806
1017 LysSerMetValHisHisLeuTyrPheAsnAspLysLeuGluAlaProAr 1033
 :::||||| :::||||| :::||||| :::||||| :::
2807 AAGTCAATGGCCACATACTTGTGGAGAGAATAAGATTAAAGCTTCCCGC 2856
1033 gLysSerArgPheProGluArgPheMetAspAspIleAlaAlaLeuValS 1050
 :::||||| ::: ::: :::||||| ::: :::
2857 AGGCCAGAGATTTCCCGAGACATATCATGTCTTACATTCACCTGCTTC 2906

1050 erThrIleAlaSerAspIleValSerArgPheGlnLysAspThrGluMet 1066
 :::||||| ::: :::||||| ::: :::
2907 TTGCATATTAATTCGCCAATGTGACTATTCGGTATGGCGGAGATTCGCCGATGAG 2956
1067 ValGluArgLeuAsnThrSerLeuAlaPhePheLeuAsnAspLeuLeuSe 1083
 :::||||| :::||||| :::||||| :::||||| :::
2957 TCCAGAATATGTGACSTATAGTTGGCTAGCTTCCTGAAGCGCTGTTTGAC 3006
1083 rValMetAspArgGlyPheValPheSerLeuIleLysSerCysTyrLysG 1100
 :::||||| :::||||| :::||||| :::||||| :::
3007 ACTAATGGATGAGGATTTATTTTCAATTATTAATGACTATATATCTG 3056
1100 InValSerSerLysLeuTyrSerLeuProAsnProSerValLeuValSer 1116
 :::||||| :::||||| :::||||| :::||||| :::
3057 GATTCAGCCCCAAA.....GATCCTAAGGTTCTGGCTGAA 3091
1117 LeuArgLeuAspPheLeuArgIleIleCysSerHisGluHisTyrValTh 1133
 :::||||| :::||||| :::||||| :::||||| :::
3092 TACAAGTTTGAATTTCTGCAACAACATTTGCAATGCAGAACATTTACATTC 3141
1133 rLeuAsnLeuProCysSerLeuLeuThrProProAlaSerProSerProS 1150
 :::||||| ::: :::||||| ::: :::
3142 TCTGAACCTTGCCAATGGCATTTTGCAAAACCTTAACCTGCAG..... 3181
1150 erValSerSerAlaThrSerGlnSerSerGlyPheSerThrAsnValGln 1166
3182CGGTTCAA 3190
1167 AspGlnLysIleAlaAsnMetPheGluLeuSerValProPheArgGlnI 1183
 :::||||| :::||||| :::||||| ::: :::
3191 GATTCAATCTTGAA.....TACAGTTTATCAGATGAGTATTGCAAGCA 3234
1183 nHisTyrLeuAlaGlyLeuValLeuThrGluLeuAlaValIleLeuAspP 1200
 :::||||| :::||||| :::||||| :::||||| :::
3235 TCACCTCTTGTTGGTGTCTACTTCTGAGGGAACCTTCCATTGCTCTTCAGG 3284
1200 roAspAlaGluGlyLeuPheGlyLeuHisLysLysValIleAsnMetVal 1216
 :::||||| ::: :::||||| ::: :::
3285 ACAATTATGAG.....ATCAGATATACAGCTATCTCTGTATTA 3322
1217 HisAsnLeuLeuSerSerHisAspSerAspProArgTyrSerAspProG 1233
 :::||||| :::||||| :::||||| ::: :::
3323 AAGAATCTTTGATTAACAACATGCATTTGACACAAAGATACAGCACAAAGAA 3372
1233 nIleLysAlaArgValAlaMetLeuTyrLeuProLeuIleGlyIleIleM 1250
 :::||||| :::||||| :::||||| :::||||| :::
3373 CCAACAAGCCAAATATGACACAATTTGACCTCCCTTGTGGACTACTTT 3422
1250 etGluThrValProGlnLeuTyrAspPheThrGluThrHisAsnGlnArg 1266
 :::||||| :::||||| :::||||| ::: :::
3423 TGGAAAATATATACAGCGCATTAGCAGGTCGAGATACCTGTATCTTGTCGA 3472
1267 GlyArgProIleCysIleAlaThrAspAspTyr..... 1277
 :::||||| :::||||| :::||||| ::: :::
3473 GCCATGCCCTAATTCTGCATCCAGAGATGAGTTTCCATGTGGCTTACTTC 3522
1277 1277
3523 ACCTGCCAATAGAGGAGTCTGAGCAGCTGACAAGACACCGCTTATGGGT 3572
1278GluSerGluSerGlySerMet 1284
 :::||||| :::||||| :::||||| ::: :::
3573 CTTTTCAAATGAGCATGCAATTAAGAGAGAGATTCACAGAGTTCCCTC 3622
1285 IleSerGlnThrValAlaMetAlaIleAlaGlyThrSerValPro..... 1299
 :::||||| :::||||| :::||||| ::: :::
3623 ATCCCAAGAA.....GGAGCAACAGGATTTCCAGATCA 3654
1300GlnLeuThrArgProGlySerPheLeuLeuThrSerT 1312
 :::||||| :::||||| :::||||| ::: :::
3655 GGGCAACACACTGGTGAATAATACCCGACAGAGATTCT.....ACAAGGA 3695
1312 hrSerGlyArgGlnHisThrThrPheSerAlaGluSerSerArgSerLeu 1328

3696 GTAGTGTATCCAGTATTAACCGCGCTGGATCAGTATGAAATCAGAAGCCCTC 3745
::||| |||::: : : : : |||||
1329 LeuIleCysLeuLeuTrpValLeuLysAsnIaaspGluThrValLeuG1 1345
|||:::| | | | : : : : | | : : | |
3746 CTGATGTGCTACCTGTATATAGTAATAATGATTTCAGAAAGATACTCTCTT 3795
1345 nLysTrpPheThrAspLeuSerValLeuGlnLeuAsnArgLeuLeuAspL 1362
: : : : : : : : | | : : | | : : : : | |
3796 AACTTACTGGAATAAGATATCACCCTCAGGAGCTCATAAACATCTCTATAC 3845
1362 euLeuTyrlLeuCysValSerCysPheGluTyrlYsgLYLysValPhe 1378
| | | | : : | | : : | | | | | | : : :
3846 TTTTAGAAGTATGCTTGTTCCTTACCTTTAGATATATGGGAAAGAAACATA 3895
1379 GluArgMetAsnSerLeuThrPheLysLys...SerLysAspMetAr 1393
| | | | : : : : : : : : | | : : | | : :
3896 GCAAGGGTGCATGATGCCCTGGCTGTCAAAACACTTCGGAATAGACCGAAA 3945
1393 galALysLeuGluGluAlaIleLeuGlySerIleGlyAlaArgGlnGluM 1410
: : : : : | | | : : : : | | : : | |
3946 ATCGAAACCATGCCCTGCTCTTCGAAACAGATCAGAGAGTAATGCAGGCC 3995
1410 etValArgArgSerArgGlyGlnLeuGluArgSerProSerGlySerAla 1426
: : : : : : : : : : : : | | : : | | : :
3996 GGCTTCAG.....CATCTTAGTAGCCCTAGAAAGTTCA 4027
1427 PheGlySerGlnGluAsnLeuArgTrpArgLysAspMetThrHisTrpAr 1443
| | | | : : : : : : : : | | : : | |
4028 TTT.....ACACTTAATCACAGTTC 4047
1443 gGlnAsnThrGluLysLeuAspLysSerArgAlaGluIleGlnHisGluA 1460
: : | | | | | | | | : : : : | | : : | |
4048 TACAACAACCTGAA.....GCAGACATTTTCCACCAGG 4079
1460 laLeuIleAspGlyAsnLeuAlaThrGluAlaAsnLeuIleLeuAsp 1476
| | | | : : : : | | | | : : : : | | : : | |
4080 CACTTCTTGAAAGCAATACAGCTACTGAAAGTTTCCCTAACAGACTAGAC 4129
1477 ThrLeuGluIleValAlaGlnThrValSerVal.....ThrG1 1489
| | | | : : : : : : | | : : : : | | : :
4130 ACCATATCATTTTTCACCTCAGTGCCTCAAGACCACACTTTTAATATATGA 4179
1489 uSerLysGluSerIleLeuGlyGlyValLeuLysValLeuLeuHisSerM 1506
: : : : : : : : : : | | : : : : | | :
4180 TGGCCATTAACCCATTAATGAAAAAAGTGTTCATATATACATCTTGCCTTTC 4229
1506 etAlaCysAsnGlnSerAlaValTyrlLeuGlnHisCysPheAlaThrGln 1522
: : : : | | | | | | | | : : : : | | : :
4230 TTAATAATGGACAATCTGAAGTGTCCGTGAACATGTATTTGCCCTCAGCTG 4279
1523 ArgAlaLeuValSerLysPheProGluLeuLeuPheGluGluGluThrG1 1539
| | | | : : : : | | | | : : : : | | : :
4280 AGAGCTTTCATCAGTAAGTTCCTTCAGCATTTTTCAAAGAGAGAGTAAA 4329
1539 uGlnCysAlaAspLeuCysLeuArgLeuLeuArgHisCysSerSerSerI 1556
: : : | | | | | : : | | : : : : | | : :
4330 CATGTGTGCTGCAATTTTGTATGAGGTTTAAAGTGTGCACATCGAAGA 4379
1556 leGlyThrIleArgSerHisProSerAlaSerLeuTyrlLeuLeuMetArg 1572
| | : : : : | | : : : : | | | | | | | |
4380 TTAGCTCAACCAAGGAATGAAGCATCTGCACCTTTGTATCTTTTGATGAGA 4429
1573 GlnAsnPheGluIle.....GlyAsnAsnPheAlaArgValLysMetG1 1587
: : | | | | | | | | : : : : | | : : | |
4430 AACCACTTTGAGTATACCAAAAGGAAACCTTTTGAGAGACACATCTACA 4479
1587 nValProMetSerLeuSerSerLeuValGlyThrSerGlnAsnPheAsnG 1604
| : : : : : : : : | | : : : : | | : : : :
4480 GATTAATAATTGCTGTAAGCCAACATGATAGCTGATGACACTAAGCGGAG 4529
1604 lUgluPheLeuArgArgSerLeuLysThrIleLeuThrTrpAlaGluGlu 1620
: : : : : : : : | | | | : : : : | | : : : :

4530 GATCAAGATTTCAGAGTCTTTTATTCATTATCAATAATTTTGCAATAAGT 4579
1621 AspLeuGluLeuArgGluThrThrPheProAspGlnValGlnAspLeuVa 1637
| | | : : : : : | | : : | | | | : : | | : : | | : :
4580 GACAGACCTATGAAAGCAACTGCCCTTCCCGCAGAAAGTCAAGACTTGAC 4629
1637 lPheAsnLeuHisMetIleLeuSerAspThrValLysMetLysGlnHisG 1654
: : : : : : : : : : | | | | : : : : | | : : | | : :
4630 CAAGAGAATCCGCACTGTCTTATGCGCACTGCCCAATGAAGGAGCATG 4679
1654 lGlnLyspProGluMetLeuIleAspLeuMetTyrlArgIleAlaLysGly 1670
: : : : | | | | | | | | | | : : | | : : | | : :
4680 AGAAAGACCCGTGAATGCTAATTGATCTCCAGTATAGCTTAGCCAAGTCC 4729
1671 TyrlGlnThrSerProGlu...ArgLeuThrTrpLeuGlnAsnMetAlaG1 1686
| | | : : : : | | | | | | | | : : : : | |
4730 TATGCAAGCACCCAGAGCTCAGGAAACCCTGGCTTGATAGCATGGCCAA 4779
1686 yLysHisSerGluArgSerAsnHisAlaGluAlaAlaGlnCysLeuValH 1703
| | | : : : : : : : : : : | | | | | | : : | |
4780 GATTCAATGTAATAATGGAGATTTTTCAGAGGCTGCAGATGTGTATGTCC 4829
1703 iSSerAlaAlaLeuValAlaGluTyrlLeuSerMetLeuGluAspArgLys 1719
| | | | | | | | | | | | : : : : | | : : : :
4830 ATGTAGCAGCTCTAGTTGCAGAGTTT.....CTTCATCGAAAAAAA 4870
1720 TyrlLeuProValGlyCysValThrPheGlnAsnIleSerSerAsnValle 1736
: : : | | | | | | : : | | : : : : | | : :
4871 TTATTTCCTAACGGAATGTTCAGCGTTCAAGAAAAATTACTCCCAATATAGA 4920
1736 uGluGluSerAlaValSerAspAspValValSerProAspGluGluGlyI 1753
| | | | : : : : | | : : : : | | : : | |
4921 TGAAGAAGAGCAATGAAGAAGATGCTGGAGTATGAT..... 4960
1753 leCysSerGlyLysTyrlPheThrGluSerGlyLeuValGlyLeuLeuGlu 1769
: : : : : : : : : : | | : : : : | | : :
4961GTCCATTATAGTGAAGAGGCTTGTGAGGTTGCTAGAA 4999
1770 GlnAlaAlaAlaSerPheSerMetAlaGlyMetTyrlGluAlaValAsnG1 1786
| | : : : : : : : : | | | | | | : : : : | |
5000 CAATGTGTGATGGCTTATGGAAGGCAGAACGTTATGAAATATATTCTGA 5049
1786 uValTyrlLysValLeuIleProIleHisGluAlaAsnArgAspAlaLysL 1803
| : : : | | : : : : : | | | | : : | | : :
5050 GATTTCCAAGTTGATCGTTCCCAATTATGAGAAACGTCGTGAGTTGAGA 5099
1803 ySLeuSerThrIleHisGlyLysLeuGlnGluAlaPheSerLysIleVal 1819
| | | | : : : : : : | | : : : : | | : : : :
5100 AACTTACTCAAGTTATAGAACTCTTCATGAGACTTACACAAAAATTCTG 5149
1820 HisGlnSerThrGlyTyrlGluArgMetPheGlyThrTyrlPheArgValG1 1836
: : : : : : : : : : | | : : : : | | : : : :
5150 GAAGTTATGCATACAAAAAAGAGACCTTTTAGGCACCTTCTTCAGAGTTGC 5199
1836 yPheTyrlGlyThrLys..PheGlyAspLeuAspGluGlnGluPheValT 1852
: | | | | | | | | : : | | : : | | : : : : | |
5200 CTTTATGCGCAATCTTTTTTTGAAGAGAGATGGAAGGAGTACTCT 5249
1852 yrLysGluProAlaIleThrLysLeuAlaGluIleSerHisArgLeuGlu 1868
| | | | | | | | : : | | | | : : | | | |
5250 ATAAAGAAACCAAGCTCACTGGCCTCTCAGAAATTTCCCTTGAGACTGTGT 5299
1869 GlyPheTyrlGlyLuarGpPheGlyGluAspValValGluValIleLysAs 1885
: : : | | | | | | : : | | : : | | : : : : | |
5300 AAACCTTATGTTGTAAGATTGTTGATACGGAATGTCAAAATATATCAGGA 5349
1885 pSerAsnProValAspLysCysLysLeuAspProAsnLysAlaTyrlIleG 1902
| | | : : : | | : : : : | | | | | | : : | | : :
5350 TTCAGACAAGGTAAATGCCAAAGAGCTGATCCAAATATGCTCATATAC 5399
1902 lNleThrTyrlValGluProTyrlPheAspThrTyrlGluMetLysAspArg 1918
| : : : | | | | | | : : | | | | | | : : : : | |
5400 AAGTTACTTATGTGAAGCCTTACTTTGATGACAAAGAACTCACAGAAAGG 5449

[illegible]

seq_name: /SIDSL/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT: AAD19118

seq_documentation_block:

ID AAD19118 standard; CDNA; 7506 BP.

... AAD19118;
AC

DT 18-DEC-2001 (first entry)

DE Angiogenesis associated human thyroid regulated transcript (hTRG) cDNA.

KW Angiogenesis associated protein; AAP; cytostatic; cardiant; gene therapy;
 KW ophthalmological; vulnary; myocardial infarction; macular degeneration;
 KW diabetic retinopathy; angiogenesis; wound healing; prophylactic; vaccine;
 KW rheumatoid arthritis; psoriasis; drug screening; tumour; transplantation;
 KW cancer; therapeutic; diagnostic; human; thyroid regulated transcript;
 TRG; ss.

OS Homo sapiens.

FH	Key	Location/Qualifiers
FT	CDS	43..6366

```
FT      /*tag= a
FT      /product= "human thyroid regulated transcript (hTRG)"
```

PN W0200170808-A2.

PD 27-SEP-2001.

PF 22-MAR-2001; 2001WO-US09609.

PR 22-MAR-2000; 2000US-191134P.

XX (CURA-) CURAGEN CORP.
PA (GETH) GENENTECH INC.
XX
PI Rastelli Lk, Gerritsen M,
XX
DR WPI: 2001-602775/68.
DR P-PSDB; AAEl1889.

PT Novel angiogenesis associated polypeptides and polynucleotides encoding
PT the polypeptides, useful for modulating angiogenesis and for treating
PT tumors and cancers -

PS Claim 6; Page 18-21; 159pp; English.

CC The invention relates to angiogenesis associated proteins (AAP) and their
CC corresponding cDNA molecules, which are useful for modulating
CC angiogenesis. AAP proteins and nucleic acids are useful for promoting
CC wound healing, for example after organ transplantation, and in the
CC treatment of tumours, myocardial infarction, cancers, diabetic
CC retinopathy, macular degeneration, psoriasis and rheumatoid arthritis.
CC AAP proteins and DNA's are useful in potential prophylactic and
CC therapeutic applications implicated in a variety of disorders including
CC those related to angiogenesis, and also in diagnostic applications.
CC AAP cDNA is also useful in gene therapy. The invention also relates to
CC a method for screening a tissue sample for tumourigenic potential. AAP
CC proteins are used to screen drugs or compounds that modulate AAP activity
CC or expression as well as treating disorders characterised by insufficient
CC or excessive production of AAP or production of AAP forms that have
CC decreased or aberrant activity compared to the wild type protein, or
CC modulate biological function that involve AAP. The present cDNA sequence
CC encodes human thyroid regulated transcript (hTRG) protein which is an
CC angiogenesis associated protein (AAP) of the invention. Human TRG is
CC upregulated in the in vitro model of angiogenesis and is likely to be
CC involved in signal transduction between receptors and kinases. Modulation
CC of hTRG is useful to treat diseases related to thyroid stimulating
CC hormone (TSH) imbalance.

SQ Sequence 7506 BP; 2147 A; 1689 C; 1764 G; 1906 T; 0 other;

alignment_scores:		
Quality:	2345.50	Length: 2361
Ratio:	1.827	Gaps: 64
Percent Similarity:	54.384	Percent Identity: 27.742

alignment_block:

US-09-737-246-2 X AAD19118

Align seg 1/1 to: AAD19118 from: 1 to: 7506

2 AAGLuArgargalaphaalaglnLysileSerargThr...ValAlaAl 17

73 GCGGAGACTCGGAAGTTCACCCGGGGCTGAGTAAGCCGGGCACGGCGC 122

17 agluValArgLysGlnIleSerGlyGlnTyrSerGlySerProGlnLeu 34

123 CGAGCTGCCGCAGAGCGTGTCTGAGGTGGTGGCGGCTCCGTGCTCCTGG 172

34 eULysAsnLeuAsnIleValGlyAsnIleSerHisHisThrThrValPro 50

173 CAAAG..... 177

51 LeuThrgIuAlaValAspProValAspLeuGIuAspTYrLeuIleThrHi 67

178 .. CCAAGCTAATTGAGCCACTGACTATGA AATGTCATCGTCCAGAA 224

67 SPRLeuAlaValAspSerGlyProLeuArgAspLeuIleGluPro 84

225 GAAGACTCAGATCCTGAACGACTGTTACGGGAGATGCTGCTCTTCCCTT 274

84 roaspaspillegluvalvaltyrserproargaspccysargthrleuval 100

PR 11-APR-2000; 2000US-0196527.
 PR 11-APR-2000; 2000US-0196528.
 PR 11-APR-2000; 2000US-0547276.
 PR 13-OCT-2000; 2000US-0240503.
 PR 13-OCT-2000; 2000US-0240508.
 PR 13-OCT-2000; 2000US-0240543.
 PR 13-OCT-2000; 2000US-0240539.
 XX
 PA (ARBO-) ARBOR VITA CORP.
 XX
 PI Lu P, Garman JD, Candia AF;
 XX
 DR WPI; 2001-367865/38.
 DR P-PSDB; AAU04025.
 XX
 XX CLASP-5 polynucleotides, proteins and antibodies are used to prevent or
 PT treat a CLASP-5 mediated disease, such as an autoimmune disease e.g.
 PT Rheumatoid arthritis -
 PT
 XX
 PS Disclosure; Fig 9A; 188pp; English.

CC The sequence encodes a protein presented as AAU04025, included in the
CC specification which relates to a cadherin-like asymmetry protein,
CC CLASP-5, which is a transmembrane protein of the immune system involved
CC in the formation of the immune gateway. CLASP-5, polynucleotides encoding
CC it and an anti-CLASP-5 antibody are used to prevent or treat a CLASP-5
CC mediated disease, such as an autoimmune disease caused or exacerbated by
CC increased activity of TH1 cells. These diseases may include Addison's
CC disease, rheumatoid arthritis, dermatitis, Multiple Sclerosis, Systemic
CC Lupus Erythematosus and Autoimmune Thyroiditis, inflammatory conditions
CC (e.g ischaemia-reperfusion) and responses, Leukaemia, acquired
CC immunodeficiency syndrome, AIDS, syphilis, typhoid, Burkitt's lymphoma,
CC Grave's disease and insulin dependent diabetes mellitus. CLASP-5 is used
CC for inhibiting an immune response in a cell (T cell or B cell) by
CC interfering with the expression of a CLASP-5 gene in the cell, the
CC ability of a CLASP-5 protein to bind to another cell or the ability of a
CC CLASP-5 protein to bind to another protein. The antibody binding CLASP-5
CC is used to inhibit an immune response is a subject. The polynucleotide is
CC used to detect CLASP-5 expression in cells and for diagnosis of
CC diseases and disorders associated with aberrant expression of CLASP-5.
CC Note: The present sequence is included in the specification but is
CC not mentioned anywhere in the specification.

Sequence 2014 BP; 565 A; 469 C; 514 G; 466 T; 0 other;

```
alignment_scores:
    Quality: 2245.50      Length: 619
    Ratio: 4.098          Gaps: 2
Percent Similarity: 88.530 Percent Identity: 69.628
```

```
alignment_block:
US-09-737-246-2 x AAS07381 ..
```

Align seg 1/1 to: AAS07381 from: 1 to: 2014

1469 GluAlaAsnLeuIleIleLeuAspThrLeuGluIleValGlnThrVa 1485
:: ::||| |||
8 CAACGAGTTTAATCATCTCGATATGCAGAAACATTATCCAGCGAG 57

```

1485 lservalThrglUserLysglUserIleLeuglyglyValLeuIrySVal 1502
      |||::: :: |||:::||||:|||||||:|||:::|||||
58 CTCGGCTTGACTGTAAAGACAGCCTGCTGGAGGTCTTCTAGCGTC 107
```

```

1502 eulEuiHisserMetalacysasngInseralavaltyrleuGInHisCys 1518
      ||:::||||: ||:::||||:||||:||||: |||||
108 TGCGAATTCCTCGAAGCTGTGATCAGAGTACCACCACTACCTGACTCACTGC 157

```

1519 phealaThrGlnargalaleuValSerLysPheproGluLeuPheG1 1535
 ||||| |||||:::|||||
 158 TTTCACACACTCCGTCCTCATTCCCAAGTTTGAGACTTACTCTTGA 207

1535 uGIuGIuThrGIuGIuGlnCysAlaAspLeuCysLeuArgLeuLeuArgHisC 1552

208 AGAGGAGTGGACACAGTCTTTCGACCTATGTCACCAAGTCTGCACCACT 257

1552 ysserserSerIleGlyThrIleargSerHisProSerAlaSerLeuTyr 1568
 |||||::: :: |||||:: |||||
 258 GCAGCAGCAGCATGATGTCACCCGGAGGCCAAGCCCTGCGCCACCCCTTAC 307

1569 LeuLeuMetArgGlnAsnPheGluIleGlyAsnAsnDheAlaArgValLy 1585
 |||||
 308 CTCCTCATGAGTTCACTTTTGAGACCAACCAGTAATTTTGCACAGATAAA 357

```

1585 smetginValPrometSerLeuSerSerLeuValGlyThrSergInasp 1602
      |||||  |||||:::|||||  ::::|
358 GATGCAAGTAAACCATGTCCTGGCATCTTTGGTGGGAGACACACGACT 407

```

```

1602 heasnglupheueargArgSerLeuLysThrIleuThrTyrAla 1618
      ||||| | | | | | | | | | | : | | | | | : | | : :
408 TTATGAGAGACACCCTGGAGAAGATCCTTGAGGACAATTGTGGCCATTCA 457
```

1619 GLUGLWASPLEUGLWUARGGLUThrThrPheProaspGlnValGlnAs 1635
 ||||| :||| ||||| ||||| :|||
 458 GAAGAGGACACAGCCATGCAGATGACTCCTTTTCCACCCAGGTGGAGGA 507

```

1635 pleuvalPheasnleuHismetileleuSerAsprnVallysmetylsg 1652
      ::::: |:::::| |::::| |::::| |::::|
508 ACTTCTGTGTAATCTGAATAGCATCTTATATGACACACAGTAAATGAGGG 557
```

```

1652 luHisglnluaspProgluMetleuileaspLeumetryrrargileala 1668
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
558 AATTTCAGGAGATCCTGATGCTTATGGATCTCATGTACAGAAATTGCC 607

```

1669 LysGlyTyrGlnThrSerProGlu..ArgLeuThrTripleGlnAsnMe 1684
 |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
 608 AAGAGTACCAGCAGCATCTCTGATCTGGCGCTGACCTGGCTCCAGACAT 657

1684 tAlagIyLShISerGIuarGSeRasnHIsalaglualaalaglncysl 1701
||| |||| ::::: ||||| ::||
658 GGCAGAGAACACACCAGAGAAGTGTCTACACGAGCGTCGCATGTGCC 707

```

1701 euValHisSerAlaAlaLeuValAlaGluTyrLeuSerMetLeuGluasp 1717
|||||::|||||
708 TGGTGACGCGCGGTGGTGGCTGAGTATCTGAGCATGCTCGAGGAC 757

```

```

1718 ArgLysTyrLeuProValGlyCysValThrPheGlnAsnIleSerSerAs 1734
      :::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
758 CACAGCTACCTGCCCCGTGGCAGTGTCCAGCTTCCAGAAATATTCTTCCAA 807

```

```

1734 nValLeugluInSerAlaValSeraspValValSerProaspLug 1751
      |||||:::|||||:|||||:|||||:|||||:|||||:
808 TGTGCTGGAGAGCTGTGTCTCTGAGGACACCCCTGTACCTGACGAGG 857

```

```

1751 luglyllecysserglyltyrphethrgluserglyleuValglyleu 1767
::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
858 ATGGGTGTGCGCAGGCCAGTACTTACCGAGAGTGGCCCTGTAGGCCCTC 907

```

1768 LeuGlIaIaIaIaSerPheSerMetAlaGlyMetTyrGluAlaVa 1784
 ||||| ||||| ::|||::|||::|||
 908 CTGAGCAGGCCGCGAGCTCTTCAACACGGAGGCTTATATGACACAGT 957

```

1784 lAsnguIuValTyrLrLySValleuIleProIleHisgluAlaasnArGAspa 1801
      |||||
      |||||
      |||||
958 TAATGAGGTCTACAAGCTGTGTATCCCATCTAGAGCGCATGAGAAAT 1007
      |||||
      |||||
      |||||

```

```

1801 lalysLysLeuSerThrIleHisGlyLysLeuGlnGluAlaPheSerLys 1817
      ::|||||::: |||:::|||||:::|||||:::
1008 TCCGGAAGCTGACACTCATTACACAGCAAGCTGCAGAGAGCCCTTCGACAGC 1057

```

```

1818 ILevalHisGlnSerThrGlyTyrPgluArgMetPheGlyThrTyrPheAr 1834
      |||||:::|||||
1058 ATCGTTAACAGGAT.....CATAAGAGAATGTTGGAACTTACTTCCG 1101

```

1834 gvaIGlyPheTYrGLYThrLysPheGLYAspLeuAspGLuGLuIupheV 1851

1102 AGTTGGTTTCTTTCGATCCAAATTGGGGATTGGATGTAACAGAGTTTG 1151
1851 aLtyrLysGluProAlaIlethrLysLeuAlaGluIleSerHisArgLeu 1867
|||||
1152 TCTACAAGAGCGCTGCAATTACCAAGCTTCTGAGATCTCATAGACTA 1201
1868 GluGlyPheThrGlyGluArgPheGlyGluAspValGluValIleIlely 1884
|||||
1202 GAGGCATTTTATGGTCAATGTTTGTGTCAGAAATTGTGGAAGTGAATA 1251
1884 sAspSerAsnProValAspLysCysLysLeuAspProAsnLysAlaTyrI 1901
|||||
1252 AGACTCCACTCCTGTGGACAAACCAAGTTGGATCCTAACACAGCCCTACA 1301
1901 IeGlnIleThrTyrValGluProTyrPheAspThrTyrGluMetLysAsp 1917
|||||
1302 TACAGATCACTTTTGTGGAGCCCTACTTTGATGAGTATGAGATGAAGAAG 1351
1918 ArgIleThrTyrPheAspLysAsnTyrAsnLeuArgArgPheMetTyrCy 1934
|||||
1352 AGGGTCACATACTTTGAGAAGAAATTCAACCTCCGAGGTTCAATGACAC 1401
1934 sThrProPheThrLeuAspGlyArgAlaHisGlyGluLeuHisGluGlnP 1951
|||||
1402 CACCCCGTTCACCCCTGGAGGGCGGCCCTCGGGAGAGAGCTGCATGAGCAGT 1451
1951 heLysArgLysThrIleLeuThrThrSerHisAlaPheProTyrIleLys 1967
:::
1452 ACAGAAAGAACACAGCTCCTGACCACTATGCACGCCCTCCCTACATCAAG 1501
1968 ThrArgValAsnValThrHisLysGluGluIleIleLeuThrProIleG 1984
|||||
1502 ACCAGGATCAGCGCTCATCCAGAGAGAGAGTTGTTTGTGACACCGATTGA 1551
1984 uValAlaIleGluAspMetGlnLysLysThrGlnGluLeuAlaPheAlaT 2001
|||||
1552 AGTGGCATTTGAAGACATGAAGAGAAGACCCTGCAGTTAGCAGTTGCCA 1601
2001 hrHisGlnAspProAlaAspProLysMetLeuGlnMetValLeuGlnGly 2017
:::
1602 TTAAACGAGAGCCGCCCTGATGCAAAGATGCTTCAGATGTGCTGCAAGGC 1651
2018 SerValGlyThrThrValAsnGlnGlyProLeuGluValAlaGlnValPh 2034
|||||
1652 TCTGTGGAGCTACTGTAAATCAGGGACCACTGGAAGTAGCCCAAGTGT 1701
2034 eLeuSerGluIleProSerAspProLysLeuPheArgHisHisAsnLysL 2051
|||||
1702 TTTGGCTGAATTCCTGCTGATCCAAACTCTATCGACATCACACAAGCT 1751
2051 euArgLeuCysPheLysAspPheThrLysArgCysGluAspAlaLeuArg 2067
|||||
1752 TGAGTTATGCTTTAAGGAATTTCATCATGAGATGTGGTGAAGCTGTAGAG 1801
2068 LysAsnLysSerLeuIleGlyProValGlnLysGluTyrGlnArgGluLe 2084
|||||
1802 AAAAACAAAGCGTCTCATTCACGGCAGACCAAGAGGGAATATCAGCAGGAAC 1851
2084 uGlyLys 2086
|
1852 CAAAAAG 1858

seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA2000.DAT: AAC74524

seq_documentation_block:

ID AAC74524 standard; cDNA; 6816 BP.

AC AAC74524;

DT 08-FEB-2001 (first entry)

Human ORFX ORF79 polynucleotide sequence SEQ ID NO:157.

KW Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
KW vulnerable; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW hypotensive; dermatological; immunosuppressive; antinflammatory;
KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
KW antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW cholesterol ester storage; systemic lupus erythematosus; infection;
KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW bone damage; cartilage damage; antinflammatory disease; coagulation;
KW thrombosis; contraceptive; ss.

XX Homo sapiens.

OS W0200058473-A2.

PN 05-OCT-2000.

PD 31-MAR-2000; 2000WO-US08621.

XX 31-MAR-1999; 99US-0127607.

PR 02-APR-1999; 99US-0127636.

PR 05-APR-1999; 99US-0127728.

PR 30-MAR-2000; 2000US-0540763.

XX (CURA-) CURAGEN CORP.

PI Shimkets RA, Leach M;

XX WPI; 2000-602362/57.

DR P-PSDB; AAB40315.

XX Novel nucleic acids and peptides derived from open reading frame X,
PT useful for treating e.g. cancers, proliferative disorders,
PT neurodegenerative disorders and cardiovascular disease -
PT

XX Claim 5; Page 502-506; 5507pp; English.

CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
CC sequences have activities such as: cytostatic; hepatotropic; vulnerable;
CC antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
CC osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
CC antidiabetic; hypotensive; dermatological; immunosuppressive;
CC antinflammatory; antibacterial; antiviral; antifungal; antineumatic;
CC antithyroid; and antianaemic. The sequences can be used for determining
CC the presence of or predisposition to, or preventing or treating
CC pathological conditions associated with an ORFX-associated disorder. The
CC nucleic acids can be used to express ORFX proteins in gene therapy
CC vectors. The proteins and nucleic acids may be used to treat cancers,
CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
CC graft vs host disease, cardiovascular disease, diabetes mellitus,
CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
CC nocturnal haemoglobinuria, antinflammatory disease; to enhance
CC coagulation; to inhibit thrombosis; and as a contraceptive.

XX Sequence 6816 BP; 1977 A; 1523 C; 1557 G; 1757 T; 2 other;
SQ

alignment_scores:

Quality: 2244.50 Length: 2149
Ratio: 1.861 Gaps: 56
Percent Similarity: 56.119 Percent Identity: 28.478

alignment_block:

US-09-737-246-2 x AAC74524 ..

Align seg 1/1 to: AAC74524 from: 1 to: 6816

```

 94 ArgAspCysArgThrLeuValSerAlaValProGluGluSerGluMetAs 110
    ::||| :::: ||::: ::||| |||||
101 CAGGACAAAGTAGTTATCTCTGGCAGACAGACGTGAAGTGGAATG.. 148
110 pProHisValArgAspCysIleArgSerTyrThrGluAspTrpAlaIleV 127
    ::||| ::||| ::||| ::|||
149 .....GAAAGATGGATCACAA 164
127 allLeArgLysTyrHisLysLeuGlyThrGlyPheAsnProAsnThrLeu 143
    ::||| ::||| ::||| ::|||
165 TTCTAAATGAAGATCCTCCAGCTCAAC..... 190
144 AspLysGlnLysGluArgGlnLysGlyLeuProLysGlnValPheGluSe 160
    ::||| ::||| ::||| ::|||
191 .....TTGCAAGC 198
160 r.....AspGluAlaProAspGlyAsnSerTyrGlnAspGlnAspa 175
    ::||| ::||| ::||| ::|||
199 TGCATATGCAGAAAGCGAATGCGGACTCTCAGAGATGATGAACAA. 247
175 sPLeuLysArgArgSerMetSerIleAspAspThrProArgGlySerTrp 191
247 ..... 247
192 AlaCysSerIlePheAspLeuLysAsnSerLeuProAspAlaLeuLeuPr 208
    ||| ::::: ::::: ||::: |||||
248 .....AGCAAATGGAAAGGTTCTGTCGCGTTTAGATAGCTACCTGCC 291
208 oAsnLeuLeuAspArgThrProAsnGluGluIleAspArgGlnAsnAspa 225
    |::||| ::::: |||||
292 GGAAGCTTGCCAAAGAGTGCAGAGAGAGCAAGCAAAATC..... 325
225 spGlnArgLysSerAsnArgHisLysGluLeuPheAlaLeuHisProSer 241
    :::: ||||| ::::: ||||| ||| ::::
326 ..AAACTGAAAAGTGAAGCAGAGTCAAACTTTTATTG....GAC 367
242 ProAspGluGluProIleGluArgLeuSerValProAspIlePro.. 257
    ||||| ::::: ::::: ||::: |||||
368 CCAGATGCCCAGAGCTTGACTTCTCATCAGCTGAGCCAGAAAGTGAAGTC 417
258 ...LysGluHisPheGlyGlnArgLeuValLysCysLeuSerLeuL 273
    ::||| ||||| ::||| ::||| ||||| ::|||
418 ATTTGAAGAGAAGTTGGAAAAAGGATCTTGTCAAGTGAATGATTTAT 467
273 yspheGluIleGlu..... 277
    ::||| :::::
468 CTTTCAATTTGCAATGCTGTGTGGCCGAAATGAAGAAGACCACACTACA 517
278 ..IleGluProIlePheAlaSerLeuAlaLeuTyrAspValLysGluLy 293
    ::||| ::||| ::||| ::||| ::||| ::|||
518 AATGTTGAACCTTCTTTGTACTCTATCCCTGTTTGACATAAATACAA 567
293 sLysLysIleSerGluAsnPheTyrPheAspLeuAsnSerGluGlnMetL 310
    ::||| ::||| ::||| ::||| ::|||
568 CCGGAAGATTTCTGCCGATTTCCACGTAGACCTGAACCATTTCTCAGTGA 617
310 ysgLyLeuLeuArgProHisValProProAla..... 320
    :::: ||| ||||| |||
618 GGCATAATGAT.CGCCACACAGTCCCGCGCGTGATGAATGGCAGTGGGCC 666
321 .....AlaIleThrThrIle 325
    |||
667 GAAACCCATCTGCCCTCAGGGGCATCCTTCATGAAGCCGCATGACAGTA 716
325 uAlaArgSerAlaIlePheSerIleThrTyrProSerGlnAspValPheL 342
    ::::: ||||| ::||| ||| ||::: |||||
717 TCCGAAGCAGGGAATATTTCAGTCACTGTCTCATCCAGATATATTTC 766
342 euValIleLysLeuGluLysValLeuGlnGlnLysAspIleGlyGluCys 358
    ||||| ::::: ||||| ||||| ||||| ::|||
```

```

767 TTGTGGCCAGAAATTGAAAAAGTCCTT..CAGGGGAGCATCCACACTTGC 813
359 AlaGluProTyrMetIlePheLysGluAlaAspAlaThrLysAsnLysG 375
    ||||| ||||| ||::: ||::: ||::: ||:::
814 GCTGAGCCATATATAG.....AAAAGTTCAGACTCTTCTAAGGTGGCCCA 857
375 uLysLeuGluLysLeuLysSerGlnAlaAspGlnPheCysGlnArgLeuG 392
    ::||| ::||| ::||| ||| ||||| |||||
858 GAAGGTGCTGAAG.....AATGCCAAGCAGGCATGCCAAAGACTAG 898
392 lYlYsTyrArgMetProPheAlaTrpThrAlaIleHisLeuMetAsnIle 408
    ||::: ||||| ||||| ||||| ::|||
899 GACAGTATAGAATGCCATTGCTTGGCAGCAGACATTTGTTAAG... 945
409 valSerSerAlaGlySerLeuGluArgAspSerThrGluValGluIleSe 425
    ::::: ||||| ::||| ::|||
946 ..GATGCATCTGGAATCTTGACAAAAATGCC..... 975
425 rThrGlyGluArgLysGlySerTrpSerGluArgArgAsnSerSerIleV 442
    ||| ||::: |||||
976 .....AGATTTTCTGCCATC. 990
442 alGlyArgArgSerLeuGluArgThrThrSerGlyAspAspAlaCysasn 458
990 ..... 990
459 LeuThrSerPheArgProAlaThrLeuThrValThrAsnPhePheLysG 475
    ::::: |||
991 .....TACAGGCA 998
475, nGluGlyAspArgLeuSerAspGluAspLeuTyrLysPheLeuAlaAspM 492
    |::: ||||| ::||| ::||| ::||| ::|||
999 AGACAGCAATAAGCTATCCATGATGACATGCTCAAGTTACTTGACAGACT 1048
492 eLArgArgProSerSerValLeuArgArgLeuArgProIleThrAlaGln 508
    ::||| ::||| ::||| ||| ::|||
1049 TTGGGAAACCTGAGAAG..ATGGCTAAGCTCCAGTGATTTTAGGCAAT 1095
509 LeuLysIleAspIle..SerProAlaProGluAsnProHisTyrCysLe 524
    ||| ||| ||| ::::: ||::: ||::: |||
1096 CTAGACATTACAAATTGATTAATGTTTCTCAGACTTCCCTAATTATGTTAA 1145
524 uThrProGluLeuLeuGlnValLysLeuTyrProAspSerArgValArgp 541
    :::::
1146 TTCATCATATACATTCCACAAAACAATTGAAACCTGCAGTAAAACTCCA 1195
541 roThrArgGluIleLeuGluPhe..ProAlaArgAspValTyrVal... 555
    ||| ||::: ||||| ||:::
1196 TCACGTTTGAAGTGGAGGAATTTGTGCCCTGCATACCAAAACACACTCAG 1245
556 ProAsnThrThrTyrArgAsnLeuLeuTyrIleTyrProGlnSerLeuAs 572
    ||| ||| ||| ||| ||||| ::||| ::|||
1246 CCTTACACCATCTACACCAATCACCTTTACGTTTATCTTAAGTACTTGAA 1295
572 nPheAlaAsnArgGln.....GlySerAlaArgAsnIleThrValLysV 587
    :::::
1296 ATACGACAGTCAGAAAGTCTTTTGGCCAAGGCTAGAAATATTGCGATTGCA 1345
587 alGlnPheMet..TyrGlyLysAspProSerAsnAlaMetProValIle 602
    ::||| ||::: ||::: ||::: |||
1346 TTGAATTCAAAGAATTACAGATGAGGAAGACTCTCAGCCCTTAAGTGCATTT 1395
603 PheGlyLysSerSerCysSerGluPheSerLysGluAlaTyrThrAlaVa 619
    ::||| ::::: ||||| ::||| ::|||
1396 TATGGCAGACCTGCTGGGCCAGCTTTTCACAAGAAGCGCCTTGTGCTGCACT 1445
619 lValTyrHisAsnArgSerProAspPheHisGluGluIleLysValLysL 636
    |::: ||||| ::||| ::||| ::||| ::|||
1446 TTTTACACCATCACCAAAACCCAGAAATTTTATGATGAGATTAAATAGAGT 1495
636 euProAlaThrLeuThrAspHisHisLysLeuLeuPheThrPheTyrHis 652
    ||||| ||| ::::: ||||| ||||| ::|||
1496 TGCCCACTCAGCTGCATGAAGAACCAACCTGTGCTGCATCTTCCAT 1545
```




293 slYslySIleSerglUasnpheTYrPheaspLeu.....AsnSerg 307
::: ::::::::::: ::|||::: ::|||:
365 GTTTCGAAGTGTGCGGCACATGATCTATGACCTTATTGAATGGCGATCAC 414
lucIlnmetlYsglYleuLeuArgProHisValProProAlaAlaIleThr 323
::: :::::||| ||| ||| :::
415 AAATCTCTTCTGAACTCTG.....CCTCAGATGAACTCAAA 452
324 ThrleuAlaArgSerAlaIlePheSerIleThrTYrProSerglInaspVa 340
||| ::::::::::: ::||| ||| :::::
453 GAACCTGAAGAAGAGGTTCACAGCCCAAAATTGATTATGGAACACA.... 497
340 lPheleuValIlelYsleuGlulYsValleuGlnGlnlYaspIleGlyG 357
::: ::|||::: |||::: ::::: ||| |||:
498ATTCTAGATTTTGGACCTGGTGTAGAGATGAAGAT..GGGA 537
357 lUcYsAlaGlUProTYrMet.....IlePheLYs..... 366
::: ::|||::: :::
538 ATATTTTGGATCCAGAAATTAACTAGCAGCATTAAGTCTCTTCAGAGCTCAT 587
367 GlUAlaaspAlaThrLYsasnYsglULysleuGlULysleuLYsSergl 383
||| ::::: ::::: |||::: |||::: |||::: |||::: |||::: |||:::
588 GAAATAGCTTCTAAACAAGTGGAGAAAAGGTTACAAGAGAAAATCTCA 637
383 nAlaaspGlnPhe..CysGlnArgleuGlyLYsTYr..ArgMetProP 398
| ::::: ::::: ::|||::: ::|||::: |||
638 AAAGCAGACATAGATATTAAACAGACAAGCCAAAGTTTGCTGCAACCCCTT 687
398 heaIaTrpThrAlaIleHisIleuMetAsnIleValSerSerAlaGlySer 414
::: ::::: ||| |||::: |||::: |||
688 CTCTGGCCTTGTGTGAACTCTCAAAATGTTGTTGTAATAATAGCA... 734
415 leuGlUArgaspSerThrGlUValGluIleSerThrGlyGlUArgLYsGl 431
::: ::::: ||| |||::: |||::: :::
735GAAGATGCTGAAGTCTCATGTCTCTATATGACCCCTGTGCA 775
431 ySer.....TrpSerGlU..... 435
||| ||| |||::: |||::: |||::: |||::: |||::: |||:::
776 GTCCCAATTCTATCAGTGAGAACTACTGCTGCTGCTGCTCAGTTTCAGGAT 825
436ArgArgAsnSerSerIle 441
||| ||| :::
826 TACCTAAAGACATAGACAGATTACATAATTGCGAGCCGTTTACTGAC 875
442 ValGlyArgArgSerleuGlUArg..... 449
::: ||| ::::: |||::: |||::: |||::: |||::: |||::: |||:::
876 CTCGGAAAGCAAGACCTGAAAAGGAGAAAATCAGTTTGTCTGTCTAGAT 925
450ThrThrSerGlyaspaspAlaCysasnLeuT 460
::: ::::: ||| |||::: |||::: |||::: |||::: |||::: |||:::
926 TGTTCGCGTGGTGCATGAGCTGAGGGAGACAACAACACAGAAACTGA 975
460 hrSer.....PheArgProAlaThrleuThrValThrAsn 471
||| ||| |||::: |||::: |||::: |||::: |||::: |||::: |||:::
976 CCTCGGGGTTGCGGCGCAGCTTTTGAGATGGCTGTGATGATGTACACAGAT 1025
472 phePheLYsGlnGlUArgleuSeraspGluaspLeuTYrLYsPh 488
::: ::::: ::::: |||::: |||::: |||::: |||::: |||::: |||:::
1026 ATAATA.....AATGAAAAGTAGATGATGAAGATTAAGCAGCATTT 1066
488 eleuAlaaspMetArgArgProSerSerValleuArgArgleuArgProI 505
|::: ::::: ||| |||::: |||::: |||::: |||::: |||::: |||:::
1067 CATTCCC.....TTTCAGCCGG 1083
505 leThrAla.....GlnleuLYsIleaspIleSerProAla 516
::: ::::: ||| |||::: |||::: |||::: |||::: |||::: |||:::
1084 TGGCAGGGGAGATGACTTCTTCAGACTGTTATTAACAAAGTCATCGCT 1133
517 ProGluasnProHisTYrCysleuThrProGluLeuLeuGlnValLYsLe 533
::: ::::: ||| |||::: |||::: |||::: |||::: |||::: |||:::
1134 GCCAAAGAAAGTCAACCACCAAGGGGCAAGGTTTGTGGGTAACATTTGAAATT 1183
533 uTYrProaspSerArgValArgProThrArgGluIleleuGlnPhePro. 549

```

1184 ACTTCCTGGAGAT..ATCCATCAGATCCGAAAA....GAGTTTCCGC 1224
550 .....AlaArg.....Asp 552
1225 AATTAGTGGACAGGACGACCACAGCTGTGGCTCGAAAAACAGGTTTCCGGAG 1274
553 ValTyrValProAsnThrThrTyrArgAsnLeuLeuTyrIleTyrProG1 569
1275 AATAATCATGCTGCTGTGATGTT..CGAAATGATATATCATGTATGTAACATTAGT 1321
569 nSerLeuAsnPhe..AlaAsnArgGlnGlySerAlaArgAsnIleThrV 585
1322 TCAAGGAGATTTTGATATAAGGAGCAAAACAACACAGCGAAGACGTTGAGG 1371
585 aLysValGlnPheMetTyrGlyGluAspProSerAsnAlaMetProVal 601
1372 TCACGGTGTCT..GTGTACGATGAGGATGGGAAACGATTAGACATGTG 1418
602 IlePhe.....GlyLysSerSerCysSerGluPheSerLysGluAl 615
1419 AATTTCCTCCGGGTGCTGTGTGATGAGGACGATTTTCAGAG..... 1454
615 aTyrThrAlaValValTyrHisAsn..ArgSerProAspPheHisGluG 631
1455 .TACAATCTGTGATTTACTACCAAGTAAAGCAGACCGCTGTTTGAGA 1503
631 LuIleLysValLysLeuProAlaThrLeuThrAspHisHisLeuLeu 647
1504 CTGTTAAGGTGGCCATTTCCCATGACGAGGACGTTAACCCGACGTCACCTTCGG 1553
648 PheThrPheTyrHisValSerCysGln..GlnLysGlnAsnThrProLe 663
1554 TTGTACCTTCCGCCACAGTCAATCAGACAGACTCTAAGGATTAATCTGAGAA 1603
663 uGluThrProValGlyTyrThrTyrIlePheProMetLeuGlnAsnGlyArgL 680
1604 AATATTGTGCACCTAGCATTTGTCAAGCTGATGAGATACGATGTAACCAACC 1653
680 euLysThrGlyGlnPheCysLeuProVal.....SerLeuGluLys 693
1654 TGGGAGACGGAGAGACACGATCTTATCTGTATAGGCCCGAAGCAAGAAG 1703
694 ProProGlnAlaTyrSerValLeuSerProGluValProLeuProGlyMe 710
1704 CTGGAAGATGCTGCCACGTACTTGAGT.....CTGCCCTCCAC 1741
710 LuysTTrValAspAsnHisLysGly..... 718
1742 GAAGCAGAGATTGGAAGAAAGGCCCACTCGGCCACCGCAAGACATGC 1791
719 .....ValPheAsnValGluVal 724
1792 AGAGCCTTGGGAGCTGCACCACTTAAGCAAGGACTCCTTCCAGATCTCCACG 1841
725 ValAlaValSerSerIleHisThrGlnAspProTyrLeuAspLysPhePh 741
1842 CTGTGTGTCTCCACCAACTGACTCAGAACGTGACCTT..... 1880
741 eaLalaLeuValAsnAlaLeuAspGlnHisLeuPheProValArgIleGlyA 758
1881 .....CTGGGGC 1887
758 spMetArgIleMetGluAsnAsnLeuGluAsnGluLeuLysSerSerIle 774
1888 TCTTGAATGGCGCTCCAACACC....AGCCTGCTGCACAGCAACTTG 1931
775 SerAlaLeuAsnSerSerGlnLeuGluProValValArgPheLeuHisLe 791
1932 AGGCAGCTGATGAAAGTGCATGTGTGTAAGTAGTAGAAGTTCTTCCAGGA 1981
791 uLeuLeuAspLysLeuIleLeuLeuValIleArgProProValIleAlaG 808

```



```
4354 GGAAGGAGAGAAAAACCCAGACATGTAATTGGCAATATGTGATCGAG 4403
1954 LysThrIleLeuThrThrSerHisAlaPheProTyrIleLysThrArgVa 1970
      ::::|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
4404 AGAACCATATATACAACTGCATATATAATTACCTGGAATTTTAAGGTGCTT 4453
1970 lasnValThrHisLysGluGluIleIleLeuThrProIleGluValAlaI 1987
      ::::|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
4454 TGAGGTCAAGTCTGTTTTCATGTGTGAAATACAGCCCTGGAGAAATGCCA 4503
1987 IeGluAspMetGlnLysLysThrGlnGluLeuAlaPheAlaThrHisGln 2003
      |||||  |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
4504 TCGAGACCATGCAGCTGCAGAACAGACAAGATCAACAGCATGGTGCAGCAG 4553
2004 AspProAlaAspProLysMet.....LeuGlnMetValLeuGln 2016
      |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
4554 CACCTGGATGACCCCAAGCTGCCCCATCAACCCGCTCTCCATGCTCCTGAA 4603
2016 nglySerValGlyThrThrValAsnGlnGlyProLeuGluValAlaGlnV 2033
      :|||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
4604 CGGCATCGTGGACCCAGCTGTCTATGGGGGCTTCGCAAACTACGAAAGG 4653
2033 alpheLeuSer.....GluIleProSerAspProLy 2043
      ::|||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
4654 CCTTCTTTACAGACCGGTACCTGCAGAGACACCTGAGGCCCATGAAGAAG 4703
2037 .....GluIleProSerAspProLy 2043
      :  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
4704 ATCGAGAAGCTCAAGACCTGATTGCTTGGCAGATTCTTCTTGCCCGCA 4753
2043 sleuPheArg..HisHisAsnLysLeuArgLeuGlyCysPheLysAspPheT 2059
      :  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
4754 AGGGATCAGAAATCCATGGAGACAAAGTCAACGAGGACACTGAGGCCGTTC 4803
2059 hrlYsArgCysGluAspAlaLeuArgLysAsnLysSerLeuIleGlyPro 2075
      ::|||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
4804 ACGAGAGGATGGAGGCCCTGTTCAACACAGCTGAAGGAA.....AAG 4844
2076 ValGlnLysGluTyr 2080
      |||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
4845 GTGGAGAAAGAGTAC 4859

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-971-988-1

seq_documentation_block:
; Sequence 1, Application US/08971988
; Patent No. 5786461
; GENERAL INFORMATION:
; APPLICANT: Michiyuki MATSUDA et al.
; TITLE OF INVENTION: CDNA OF DOCK180 GENE AND DOCK180 PROTEIN
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 500 kb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/971,988
; FILING DATE: 17-NOV-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/588,985
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek, Jr.
; REGISTRATION NUMBER: 33,367
```

```
REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX:
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6519 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL:
; ANTI-SENSE:
; FRAGMENT TYPE:
; ORIGINAL SOURCE: spleen cell of homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 24..5619
; US-08-971-988-1

alignment_scores:
      Quality: 242.50      Length: 2005
      Ratio: 0.287      Gaps: 101
Percent Similarity: 42.195      Percent Identity: 19.352

alignment_block:
US-09-737-246-2 x US-08-971-988-1 ..

Align seg 1/1 to: US-08-971-988-1 from: 1 to: 6519

277 GluIleGluProIlePheAlaSerLeuAlaLeuTyrAspValLysGluLy 293
      |||  ::  |||  |  |  |  |  |  |  |  |  |  |  |  |  |  |
315 GAGTGGTCCACCATCTGGAGGCGACGCTTACGTGCAAGATTAACAGGAGAT 364
293 sLysLysIleSerGluAsnPheTyrPheAspleu.....AsnSerg 307
      ::  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
365 GTTTCGAAGTGTGCGGCACATGATCTATGACCTTATTGAATGGCGATCAC 414
307 IuGlnMetLysGlyLeuLeuArgProHisValProProAlaAlaIleThr 323
      ::  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
415 AAATTCTTCTGGAAGTCTG.....CCTCAGAGATGAAGTCAAA 452
324 ThrLeuAlaArgSerAlaIlePheSerIleThrTyrProSergLAspVa 340
      |||  ::::  |  |  |  |  |  |  |  |  |  |  |  |  |  |
453 GAACCTGAAGAGAAGGTCAACAGCCCAAAATTGATTATGAAACAGA..... 497
340 IPhelLeuValIleLysLeuGluLysValLeuGlnGlnGlyAspIleGly 357
      ::::  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
498 .....ATTCTAGATTGGACCTGTGTTAGAGATGACAT...GGGA 537
357 IuCysAlaGluProTyrMet.....IlePheLys..... 366
      ::  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
538 ATATTTTGGATCCAGAAATTAAGTACGACGATTAGTCTCTTCAGAGCTCAT 587
367 GluAlaAspAlaThrLysAsnLysGluLysLeuGluLysLeuLysSergI 383
      |||  ::  ::  |  |  |  |  |  |  |  |  |  |  |  |  |  |
588 GAAATAGCTTCTAAACAAGTGGAGGAAAAGTTTACAAAGAGAAAAATCTCA 637
383 nAlaAspGlnPhe...CysGlnArgLeuGlyLysTyr...ArgMetProp 398
      |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
638 AAAGCAGACATAGATATTTAACAGACAAAGCAAGTTTGCATCAACCCCTT 687
398 heAlaTrpThrAlaIleHisLeuMetAsnIleValSerSerAlaGlySer 414
      ::::  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
688 CTCTGGCCTTGTGTGACCTCAAAAATGTGTTGTGAATAATAGA... 734
415 LeuGluArgAspSerThrGluValGluIleSerThrGlyGluArgLysGln 431
      ::::  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
735 .....GAAGATGCTGAAGTCTCATGTCTCTATATGACCCCTGTGGA 775
431 ySer.....TrpSergIu..... 435
```

```
776  |||  |||||:::
    GTCCAAATTCATCAGTGAGACTACCTGTTCCGCTGCTGCCAGTTCAGGAT  825
436  .....ArgArgAsnSerSerIle  441
    TACCTAAGACATAGACAGATTACATATAATTGCCAGCCGCTGTTACTGCAC  875
442  ValGlyArgArgSerLeuGluArg.....  449
    ::|||  ::|||::|||::|||
    CTCGGAAGCAAAGACCTGAAAAGGAGAAAATCAGTTTGTCTGTCCAGAT  925
450  .....ThrThrSerGlyAspAspAlaCysAsnLeuT  460
    TGTTCGCGTGGGTCCGATGGAGCTGAGGACAACAACACAGAAACTGA  975
460  hrSer.....PheArgProAlaThrLeuThrValThrAsn  471
    |||||  |||  |||:::  |||||:::
    CCTCGGGGTGGCGCGACCTTTTGGAGTGGCTGTGATGGATGTAAACAGAT  1025
472  PhePheLysGlnGluGlyAspArgLeuSerAspGluAspLeuTyrLysPh  488
    ::::  ::  :::::|||||||
    ATAATA.....AATGAAAAGTATGATGATGAAGATTAAGCACCATTT  1066
488  eLeuAlaAspMetArgArgProSerSerValLeuArgArgLeuArgProI  505
    |::  |::  :::::|::|
    CATTCCC.....TTTCAGCCCG  1083
505  leThrAla.....GlnLeuLysIleAspIleSerProAla  516
    :::::  ||  |||::  ||
    TGGCAGGGGAGATGACTTCCTTCAGACTGTATTAACAAAGTCATCGCT  1133
517  ProGluAsnProHisTyrCysLeuThrProGluLeuGlnValLysLe  533
    ::::  ::::  ::  ::|||
    GCCAAAGAAAGTCAACCAACAGGCGGCGTTGTGGTAACATTGAATTT  1183
533  uTyrProAspSerArgValArgProThrArgGluIleLeuGluPhePro.  549
    |  |||  ::  ::::  |||::  |||||
    ACTTCCCTGAGAT..ATCCATCAGATCCGAAAA....GAGTTTCCGC  1224
550  .....AlaArg.....Asp  552
    |::  |||||
    ATTTAGTGCAGACGACACACAGCTGTGGCTCGAAAAACAGGTTTCCGGAG  1274
553  ValTyrValProAsnThrThrTyrArgAsnLeuLeuTyrIleTyrProG  569
    ::  ::|||:::  ::  |||||  ::|||:::
    ATAAATCATGCCGTGGTATGTT..CGAAATGATATCTATGTAAACATTAGT  1321
569  nSerLeuAsnPhe..AlaAsnArgGlnGlySerAlaArgAsnIleThrV  585
    ::  ::|||  ::  ::  ::|||::|||:::
    TCAAGAGATTTTGATAAAGGAAGCAAAACAACAGCGAAGACGTGGAG  1371
585  alLysValGlnPheMetTyrGlyLysProSerAsnAlaMetProVal  601
    |  |||::  ::||  |||||  ::::  |||
    TCACGGTGTCT..GTGTACGATGAGATGGGAAACGATTAGACATGTG  1418
1372  TCACGGTGTCT..GTGTACGATGAGATGGGAAACGATTAGACATGTG  1418
602  IlePhe.....GlyLysSerSerCysSerGluPheSerLysGluAl  615
    |||||  ||  ::::  |||||
    ATTTTCCCCGGGTGCTGGTGTGATGAAGCGATTTCAGAG.....  1454
1419  ATTTTCCCCGGGTGCTGGTGTGATGAAGCGATTTCAGAG.....  1454
615  aTyrThrAlaValValTyrHisAsn..ArgSerProAspPheHisGluG  631
    ||  ::|||::|||:::  ::::||  ::  ||
    TACAATACTGTGATTACTTACCAAGTAAGCAGCCACGCTGGTTTGAGA  1503
1455  TACAATACTGTGATTACTTACCAAGTAAGCAGCCACGCTGGTTTGAGA  1503
631  LuIleLysValLysLeuProAlaThrLeuThrAspHisHisLeuLeu  647
    ::|||  ::||  :::::  |||||
    CTGTTAAGGTGGCCATTCCCATCGAGACGTTAACCGCAGTCACTTCGG  1553
1504  CTGTTAAGGTGGCCATTCCCATCGAGACGTTAACCGCAGTCACTTCGG  1553
648  PheThrPheTyrHisValSerCysGln..GlnLysGlnAsnThrProLe  663
    |||||  |||  |||  ::|||:::
    TTACCTTCCGCCACACAGGTCAACAGGACTCTAAGGATAAATCTGAGAA  1603
1554  TTACCTTCCGCCACACAGGTCAACAGGACTCTAAGGATAAATCTGAGAA  1603
663  uGluThrProValGlyTyrThrTyrIleProMetLeuGlnAsnGlyArgL  680
    :::::  ::  ::::  |
```

```
1604  AATATTTCACCTAGCAATTGTCTCAAGCTGATGAGATACGATGTTACCA  1653
680  eulYsThrGlyGlnPheCysLeuProVal.....SerLeuGluLys  693
    ||::  ||::  |||  |||  ::  ::|||
    TGCAGAGACGAGACGACGATCTTATCGTCTATAAGCCGGAAGCGAAGA  1703
1654  TGCAGAGACGAGACGACGATCTTATCGTCTATAAGCCGGAAGCGAAGA  1703
694  ProProGlnAlaTyrSerValLeuSerProGluValProLeuProGlyMe  710
    ::|||  ::  |||||
    CTGGAAGATGCTGCCACGTAAGT.....CTGCCCTCCAC  1741
1704  CTGGAAGATGCTGCCACGTAAGT.....CTGCCCTCCAC  1741
710  tLysTrpValAspAsnHisLysGly.....  718
    |||  ::::|
    GAAGGACAGAGTTGGAAGAAAAGGGCCACTCGGCCACCGGCAAGAGCAT  1791
1742  GAAGGACAGAGTTGGAAGAAAAGGGCCACTCGGCCACCGGCAAGAGCAT  1791
719  .....ValPheAsnValGluVal  724
    |:::  |||:::
    AGAGCCTTGGAGCTGCACCATTTAGCAAGGACTCTTCAGATCTCCACG  1841
1792  AGAGCCTTGGAGCTGCACCATTTAGCAAGGACTCTTCAGATCTCCACG  1841
725  ValAlaValSerSerIleHisThrGlnAspProTyrLeuAspLysPhePh  741
    ::::  ||::  |||||  ||
    CTCGTGTGCTCCACCCAAACCTGACTCAGAACGTGCACCTT.....  1880
1842  CTCGTGTGCTCCACCCAAACCTGACTCAGAACGTGCACCTT.....  1880
741  eAlaLeuValAsnAlaLeuAspGluHisLeuPheProValArgIleGlyA  758
    ::|||
    .....CTGGGC  1887
1881  .....CTGGGC  1887
758  spMetArgIleMetGluAsnAsnLeuGluAsnGluLeuLysSerSerIle  774
    ::::  ::|||::  ::  |||:::
    TCTTGAATGGCGCTCCAAACACC....AGCCTGCTGCAGCAGACACTG  1931
1888  TCTTGAATGGCGCTCCAAACACC....AGCCTGCTGCAGCAGACACTG  1931
775  SerAlaLeuAsnSerSerGlnLeuGluProValValArgPheLeuHisLe  791
    ||  ::  ::  |||||::|||:::
    AGGCAGCTGATGAAGTCGATGGTGTGAAGTATGAAGTTCTTCACGA  1981
1932  AGGCAGCTGATGAAGTCGATGGTGTGAAGTATGAAGTTCTTCACGA  1981
791  uLeuLeuAspLysLeuIleLeuLeuValIleArgProProValIleAlaG  808
    |||||  |||::  :::::
    CACGTTGATGCCCTCTTCAACATCATGATGAG.....AACT  2019
1982  CACGTTGATGCCCTCTTCAACATCATGATGAG.....AACT  2019
808  LysIleIleValAsnLeuGlyGlnAlaSerPheGluAlaMetAlaSerIle  824
    ::::  ::::  |||::|||:::
    CAGAGAGTGAGACTTTTGACACAGTTAGTCTTGATGCTCTGATATTATC  2069
2020  CAGAGAGTGAGACTTTTGACACAGTTAGTCTTGATGCTCTGATATTATC  2069
825  IleAsnArgLeuHisLysAsnLeuGluGlyAsnHisAspGlnHisGlyAr  841
    ||::  ::  ::  |||||
    ATTGGACTG.....ATTGCTGATAGAAAATTTCAGCAT..TT  2104
2070  ATTGGACTG.....ATTGCTGATAGAAAATTTCAGCAT..TT  2104
841  gasnSerLeuLeuAlaSerTyrIleHisTyrValPheArgLeuProAsnT  858
    |||  ::|||  ::|||
    TAATCCTGTTTGGAAACTTACATTAAAGAACACTTTAGT.....  2144
2105  TAATCCTGTTTGGAAACTTACATTAAAGAACACTTTAGT.....  2144
858  hrTyrProAsnSerSerSerProGlyProGlyGlyLeuGlyGlySerVal  874
    .....  2144
2144  .....  2144
875  HisTyrAlaThrMetAlaArgSerAlaValArgProAlaSerLeuAsnLe  891
    |||||::|||  ::  ::
    GCAACGTTAGCCTACACGAAGTGT.....  2168
2145  GCAACGTTAGCCTACACGAAGTGT.....  2168
891  uAsnArgSerArgSerLeuSerAsnSerAsnProAspIleSerGlyThrP  908
    ::::  |||::|||  ::::|||
    ACADAAGTGTGAAGAACTAC.....GTGACCGGT...  2198
2169  ACADAAGTGTGAAGAACTAC.....GTGACCGGT...  2198
908  roThrSerProAspAspGluValArgSerIleIleGlySerLysGlyLeu  924
    .....  2198
2198  .....  2198
925  AspArgSerAsnSerTrpValAsnThrGlyGlyProLysAlaAlaProTr  941
    .....  2198
2198  .....  2198
941  pGlySerAsnProSerProSerAlaGluSerThrGlnAlaMetAspArgS  958
    :::::  ||::  ::  ::|||
    GCTGAGAAGCCGGAGTAATGAGCAGCTGTACAAAGCCATG.....  2240
2199  GCTGAGAAGCCGGAGTAATGAGCAGCTGTACAAAGCCATG.....  2240
```


958 erCysAsnArgMetSerSerHisThrGluThrSerSerPheLeuGlnThr 974
2240 2240
975 LeuThrGlyArgLeuProThrLysLysLeuPheHisGluGluLeuAlaLe 991
2240 2240
991 uGlnTrpValValCysSerGlySerValArgGluSerAlaLeuGlnGlnA 1008
2241AAGCGCTAGAATCC. 2255
1008 laTrpPhePhePheGluLeuMetValLysSerMetVal.....HisHis 1022
2256ATCTCAAGTTTCATCGTCGCTCCAGGATCCTGTTCAATCAA 2297
1023 LeuTyrPheAsnAspLysLeuGluAlaProArgLysSerArgPheProG1 1039
2298 CTGTAATGAAAC...AAGGAGAGGCTGAC..... 2324
1039 uArgPheMetAspAspIleAlaAlaLeuValSerThrIleAlaSerAspi 1056
2325 ...TTGCTGGAATCTTTGCTGCAGCTCTTCAGGTCATC..AATGACA 2367
1056 leValSerArgPheGlnLysAspThrGluMetValGluArgLeuAsnThr 1072
2368 TGATGAGCAGCATGTCCAGACGACCGTCGCGGTGAAGGGGCGACACTG 2417
1073 Ser...LeuAlaPhePheLeuAsnAspLeuLeuSerValMetAspArgG1 1088
2418 AAATACTTACCAACGATCGTCAACGATGTGAATTTGGTTTGAT.... 2462
1088 yPheValPheSerLeuIleLysSerCysTyrLysGlnValSerSerLysL 1105
2462 2462
1105 euTyrSerLeuProAsnProSerValLeuValSerLeuArgLeuAspPhe 1121
2463CCCAAGAGCTCAGCAAAATGTTTACTGAATTC 2495
1122 LeuArgIleIleCysSerHisGluHisTyrValThrLeuAsnLeuProCy 1138
2496 ATC.....CTCAATGTTCCCAT 2512
1138 sSerLeuLeuThrProProAlaSerProSerProSerValSerSerAlaT 1155
2513 GGGCTTGCTGACC..... 2525
1155 hrSerGlnSerSerGlyPheSerThrAsnValGlnAspGlnLysIleAla 1171
2526ATCCAGAAACTCTAC 2540
1172 AsnMetPheGluLeu.....SerValProPheArgGln.....G1 1183
2541 TGCTTGATCGAAATCGTCCACAGTGACCTCTTCACACAGCATGACTGCAG 2590
1183 nHisTyrLeuAlaGlyLeuValLeuThrGluLeuAlaValIleLeuasp. 1199
2591 AGAGATCTGCTTCCCATGATGACCGATCAGCTCAAGTACCATCTGGAGA 2640
1200ProAspAlaGluGlyLeuPheGlyLeuHisLysLysValIleasn 1214
2641 GACAGAGAGACCTGGAGGCTGCTGTCAGCTGCTCAGCCACATCCTGGAG 2690
1215 MetValHisAsnLeuLeuSerSerHisAspSerAspProArgTyrSerAs 1231
2691 GTGCTGTACAGG.....AAGACGTGGGCCCAACCCAGAGCA 2728
1231 pProGlnIleLysAlaArgValAlaMetLeuTyrLeuProLeuIleGlyI 1248
2729 CGTCCAG.....A 2736

1248 leIleMetGluThrValProGlnLeuTyrAspPheThrGluThrHisAsn 1264
2737 TTATCATGGAGAAACTT.....CTCCGACCGGTGAACCGAACCCTCATTT 2780
1265 GlnArgGlyArgProIleCysIleAlaThrAspAspTyrGluSerGluSe 1281
2781 TCCATGGGACGA.....GATTCTGAA.. 2801
1281 rGlySerMetIleSerGlnThrValAlaMetAlaIleAlaGlyThrServ 1298
2802CTCATTTGGAACCTTCGTGGCTTGACATGACAGCT....ATTT 2838
1298 aLProGlnLeuThrArgProGlySerPheLeuLeuThrSerThrSergly 1314
2839 TACGACAAATGGAAGATTACCATTTATGCCCACTTGATCAAGACTTTTGGG 2888
1315 ArgGlnHisThr.....ThrPheSerAlaGluSer.....Se 1325
2889 AAAATGAGGACTGATGTGTGATGATTCTCTAATGGAAACATTCATCATGTT 2938
1325 rArgSerLeuLeu.....IleCysLeuLeuTrpValLeuLysA 1338
2939 TAAGAACCTTCATTGGAAAGAAGCTTTACCCCTTCGACTGGGTGATCATGA 2988
1338 sn.....Alaasp 1340
2989 ACATGCTGCAAAATAAGTCTTCCTCGCAGCAATTATCATGATGACAGAT 3038
1341 GluThrValLeuGlnLysTrpPheThrAspLeuSerValLeuGlnLeuAs 1357
3039ATGCTGAACAAAAAATTTCTTGATCAAGCCAACCTTGAAGCTACA 3082
1357 nArgLeuLeuAspLeuLeuTyrLeuCysValSerCysPheGluTyrLysG 1374
3083 GCTGTGAACCACTACTTTCACCTGGCTGTTGCT..... 3116
1374 lyLysLysValPheGluArgMetAsnSerLeuThrPheLysLys..Ser 1389
3117TTCCCTTACTCAAGAGTCCCTGCAACTGAGAAATTTTTCGA 3155
1390 LysAspMetArgAlaLysLeuGlnGluAlaIleLeuGlySerIleGlyAl 1406
3156 AGTGCCAAGAGAGCCAAA.....ATCCTTAACCAAGTACGGA.. 3191
1406 aArgGlnGluMetValArgArgSerArgGlyGlnLeuGluArgSerPros 1423
3192GATATGAGGAGACAGATTGGCTTTGAATC..... 3221
1423 erGlySerAlaPheGlySerGlnGlnAsnLeuArgTrpArgLysAspMet 1439
3222AGAGACATG 3230
1440 ThrHisTrpArgGlnAsnThrGluLysLeuAspLysSerArgAlaGluI1 1456
3231 TGGTAC..... 3236
1456 eGluHisGluAlaLeuIleAspGlyAsnLeuAlaThrGluAlaAsnLeuI 1473
3237AACCTTGGTCAACACACAGATAAAGT 3261
1473 leIleLeuAspThrLeuGluIleValValGlnThrValSerValThrGlu 1489
3262 TCATTCAGAAATGCTGGGCCCAATATTAAGAAATGACATTAATTCGCCGAG 3311
1490 SerLysGluSerIleLeuGlyGlyValLeuLysValLeuLeuHisSerMe 1506
3312 ACGGAG.....CTGCCAAGCCACCATCCCATCTTCTTTGATATGAT 3355
1506 tAlaCysAsnGlnSerAlaValTyrLeuGlnHisCysPheAlaThrGlnA 1523
3356 GCAGTGTGAATTCATTCGACCCGAAGCTTCCAAATGTTT....GAAA 3399
1523 rgAlaLeuValSerLysPhePro..... 1530

```
3400 ATGAGATCATCACCAAGCTGCATCATGAAGTCGAAGGAGCAGAGAGAC 3449
1531 .....GluleuLeuPhegluGlulThrGlulGlnCysAlaAspLe 1544
3450 GAACAGTACAAAGTGTATTGTGATAA..... 3476
1544 uCysLeuArgLeuLeuArgHisCysSerSerIleGlyThrIleArgS 1561
3477 .....ATCCTTCTGGAACACTGC.....AGGA 3498
1561 erHisProSerAlaSerLeuTyrLeuLeuMetArgGlnAsnPheGluIle 1577
3499 AGCACAAA.....TACCTCGCC.....AAACA 3521
1578 GlyAsnAsnPheAlaArgValLysMetGlnValPrometSerLeuSerSe 1594
3522 GGAGAAACTTTT.....GTAAAACTCGTTGTG..... 3548
1594 rLeuValGlyThrSerGlnAsnPheAsnGluLupheLeuArgArgSerL 1611
3549 .....CGCTTAA 3555
1611 eulYsThrIleLeuThrTyrAlaGluGluAspLeuGluLeuArgGluThr 1627
3556 TGGAAAGGCTTTGGATTAT.....AGAAC 3581
1628 ThrPheProAspGlnValGlnAspLeuValPheAsnLeuHisMetIleLe 1644
3582 ATCATGCACGACGACG.....AACAAAGAAAACCGCAT 3613
1644 uSerAspThrValLysMet.....LysGluHisGlnGluAsp 1657
3614 GAGCTGCACCGTCAAATGTGCTGAATTTCACAAAGAAATTGAAAGAGAA. 3662
1657 roGluMetLeuIleAspLeuMetTyrArgIleAlaLysGlyTyrGlnThr 1673
3663 ..GAATGTATATAGGTATTGTACAG..... 3689
1674 SerProGluArgLeuThrTrpLeuGlnAsnMetAlaGlyLysHisSerG 1690
3690 .....CTCTGTGACCTGCACACAGA 3709
1690 uArgSerAsnHisAlaGluAlaAlaGlnCysLeuValHisSerAlaAlaL 1707
3710 GTGTGATACTACACCGAAGCGGCTTACACTTGCTT.....C 3747
1707 euValAlaGluTyrLeuSerMetLeuGluAspArgLysTyrLeuProVal 1723
3748 TCCATGCAAAGCTTCTTAAGTGGTCGAGGAT..... 3779
1724 GlyCysValThr.....PheGlnAsnIleSerSe 1733
3780 GTGTGTGTGGCCCACTCACCCAGCGGGGAGGGGTACAGGCCACAGCGCA 3829
1733 rAsnValLeuGluGluSerAlaValSerAspAspValValSerProAsp. 1749
3830 GGGACAGCTGAAGAGCAGCTCTACACAGAAATCATCCACTACTTCGACA 3879
1750 .....GluGluGlyIleCysSerGlyLysTyrPheThrGlu 1761
3880 AAGCAAGATGTGGAGAGGAGGCCATTCCTTGGGCAAG..... 3917
1762 SerGlyLeuValGlyLeuLeuGluGlnAlaAlaAlaSerPheSerMetAl 1778
3918 .....GAGCTAGC 3925
1778 aGlyMetTyrGluAlaValAsnGluValTyrLysValLeuIleProIleH 1795
3926 CGAGCAGTATAG.....AACGAAATGTTT..... 3950
1795 IsGluAlaAsnArgAspAlaLysLysLeuSerThrIleHisGlyLysLeu 1811
```

```
3951 .....GATTATGAGCAACTCAGCGAAATTGCTGAAAAAACAG 3986
1812 GlnGluAlaPheSerLysIleValHisGlnSerThrGlyTrpGluArgMe 1828
3987 GCTCAGTTTATGAAAAACATCGTCAAAAGTGATC.....AGGCC 4024
1828 tPheGlyThrTyrPheArgValGlyPheTyrGlyThrLysPheGlyAsp. 1844
4025 CAAGCCTGACTATTTCGTGTGGCTACTACGACAGAAGGTTCCCCACAT 4074
1845 .LeuAspGluGlnGluPheValTyrLysGluProAlaIleThrLysLeu 1860
4075 TCCTGCGGGGAAAAGTTTCATTTCACGAGGGAAGAGTATGAGCCCCGG 4124
1861 AlaGluIleSerHisArgLeuGluGlyPheTyrGlyGluArgPheGlyG 1877
4125 GAAGA'TTTGAGGCTCGGCTCTTAACTCAGTTTCCAAAC..... 4163
1877 uAspValValGluValIleLysAspSerAsnProVal..AspLysCysL 1893
4164 .....GCCGAGAAAATGAACAACAACATCTCCACGAGCGCATATGTA 4206
1893 ySLeuAspProAsnLysAlaTyrIleGlnIleThrTyrValGluProTyr 1909
4207 AAACCTCTCTGCGCAG...TATATTACGTGCTTCACAGTGAAGCCCAA 4253
1910 PheAspThr.....TyrGluMetLysAspArgIleThrTy 1921
4254 CTCGATCTGCCTCCTAAGTTTCACAGGCCAGTGTCAGAGCATTTGTAAG 4303
1921 rPheAspLysAsnTyrAsnLeuArgArgPheMetTyrCysThrProPhe 1938
4304 TTTTTCACAGGGTGAACGAGCTCCAGGATTTGAATATTTCTCGGCCAATCC 4353
1938 hrLeu..AspGlyArgAlaHisGlyGluLeuHisGluGlnPheLysArg 1953
4354 GGAAGGAGAGAAAAACCCAGACAATGAATTTGCCAATATGTGATCGAG 4403
1954 LysThrIleLeuThrThrSerHisAlaPheProTyrIleLysThrArgVa 1970
4404 AGAACCATATATACAACATGCATATAAATTACCTGGAATTTTAAGTGTT 4453
1970 IAsnValThrHisLysGluGluIleIleLeuThrProIleGluValAla 1987
4454 TGAGCTCAAGTCTGTTTCATGGTGGAATCAGCCCTGGAAGATGCCA 4503
1987 IeGluAspMetGlnLysLysThrGlnGluLeuAlaPheAlaThrHisGln 2003
4504 TCGAGACCATGACGCTGACGACAGACAGACAGATCAACAGCATGGTCAGCAG 4553
2004 AspProAlaAspProLysMet.....LeuGlnMetValLeuG 2016
4554 CACCTGATGACCCCAAGCTGCCCATCAACCCGCTCTCCATGCTCCTGAA 4603
2016 nglySerValGlyThrThrValAsnGlnGlyProLeuGluValAlaGlnV 2033
4604 CGGCATCGTGCACCCAGCTGTATGGGGGCTTCGCAAACTACGAAAGG 4653
2033 alPheLeuSer..... 2036
4654 CCTTCTTTACAGACCGGTACCTGCAGAGCACCTGAGGCCCATGAAGAAG 4703
2037 .....GluIleProSerAspProLy 2043
4704 ATCGAGAAGCTCAAGACCTGATTCCTTGGCAGATTCCTTTCTGGCCGA 4753
2043 sleuPheArg..HisHisAsnLysLeuArgLeuCysPheLysAspPheT 2059
4754 AGGATCAGAAATCCATGAGACAAAGTACCGGAGGCCACTGAGGCCGTTCC 4803
2059 hrLysArgCysGluAspAlaLeuArgLysAsnLysSerLeuIleGlyPro 2075
4804 ACGAGAGGATGAGCGCTGTTTCAAAACAGCTGAAGGAA.....AAG 4844
```



```
363 etile...PheLysGluAlaAspAlaThrLys..... 372
   :: |||||:::
5168 TAACCGAATTATAAAAGAAACATTAACAGCAGAACTATTGGCTAAGCT 5217
373 .....AsnLysGluLysLeuLysLysSerGlnAlaAspG1 386
   ::|||::: |||||::: ||| |||||
5218 ATGAATCTTGATAAAATCTACTGAGACAGCCCAAGATTTCAGGCAAAAC 5267
386 n..PheCysGlnArgLeuGlyLysTyrArgMetProPheAla.TripThra 402
   ||| ||| |||::: |||||:::|||:::||| |||
5268 TGATTTGTCCAA...CTGGCCATCTATGCACAACCTATTGCAGTGATTA 5314
402 laIleHisLeuMetAsnIleValSerSerAlaGlySerLeuGluArgAsp 418
   ||| ::| ||| ::||| |||
5315 ACATCTCACCCTCAATTGAACATCTCTCCACAAGGCATT..... 5352
419 SerThrGluValGluIleSerThrGlyGluArgLysGlySerTrpSerG1 435
   ||||| ::::: |||::: |||::: |||::: |||:::
5353 TCCACACTGACTCAAGTATTGACCGCAGAACCTCCCGTAACTATACCCA 5402
435 uArgArgAsnSerSerIleValGlyArgSerLeuGluArgThrThrs 452
   : ::|||::: |||::: |||::: |||:::
5403 ATGGAAAACGCCGCTGCATATTAAACGCCGCTGCACACCCAAAGA 5452
452 erGlyAspAspAlaCysAsnLeuThrSerPheArgProAlaThrLeuThr 468
   ::::: |||::: |||::: |||::: |||:::
5453 CTAAAGCCCTACATGCGTTCTTGATGAGTCTCGCAGTGCCTTAAGC 5502
469 ValThrAsnPhePheLys..GlnGluGlyAspArg..... 479
   ::| ::::: |||::: |||||
5503 ACATACTATATTATTCTCATACCAAAAGATCGAGAAGCAAGAAAACA 5552
480 .....LeuSerAspGluAspLeuTyrLysPheLeuAlaAspMetA 493
   ::| ::||| |||||::: |||
5553 TACAGTAATTAAAAACCGTGATGATCTATATCAATACCTATTGATCGATA 5602
493 rGArgProSerSerValLeuArg....ArgLeuArgProIleThrAla 507
   ::::: |||::: |||::: |||::: |||
5603 ACCAAGTTCCGCCGACATTAACACTACAGAGATCGCTGAAGCTATCGCT 5652
508 GlnLeuLysIleAspIleSerProAlaProGlu..... 518
   ::::: |||::: |||::: |||
5653 AGTATCCAAGTGTATATTAAACCGCGCTTGAAAAATATGAGCGAGATAC 5702
519 .....A 519
5703 TGTCAACAAGTGTACACAGCCGCTCATTTCTCACCAACTGGGATAATACA 5752
519 snProHisTyrCysLeuThrProGluLeuLeuGlnValLysLeuTyrPro 535
   || ::||| ::| ::::: |||||
5753 ATAAACGCTACAGCACTTGGCGCCGCTATGGCTAAACTCCTTTACTATCCA 5802
536 AspSerArgValArgProThrArgGluIleLeuGluPheProAlaArgAs 552
   ::::: |||::: ||||| ::||| ::|
5803 GAGAATTACATCGATCCGACGCTACGTATTGGGCAGACAAAAATGATGGA 5852
552 pValTyrValProAsnThrThrTyrArgAsnLeu..... 563
   |::: ||| ::| ::|||
5853 TACGTTGCTGCAATCCATCAGCAAGCCAATTAAATATGATACCGTAG 5902
563 ..... 563
5903 AAGATGCTTTAAATCTTACCTACATCATTCGAACAGGTGGCTAATCTG 5952
564 .....LeuTyrIleTyrProGlnSerLeuAsnPheAlaAsnArgGlnG1 578
   ||| ||| ::::: ||| ||| |||||
5953 GAAATCCTCAGCGCTACCATGACACATTAAT.....AATGATCAAGG 5996
578 Y.....SerAlaArgAsnIleThrValLysValGlnPheMetT 591
   | ::|||::: ||| |||:::
5997 ATTAACCTTACTTTATCGGACGTAGTAAAAACAGAAAGTGAATCAATATT 6046
591 Yr.....GlyGluAspProSerAsn 597
```

```

   :: |||::: |||::: |||
6047 GCGCGAGTGTGATCAACAATAATCCAGCGAAGTAAATTCCCCGCTAAT 6096
598 AlaMet..... 599
6097 GCCTGAGTGAAGTGTTACAATAATTGATTGTCCAATTAAACCCTACAAGA 6146
600 .....ProValIlePhe..... 603
6147 TACTATTACCCCCGTAATTTTCCAATCTCGCCTGTATCTTATCTGGCTGG 6196
604 ....GlyLysSerSerCysSerGluPheSerLysGluAla..... 615
6197 AACAAAAAAGCGCACTAAACAGGAAGGTGATTAACCCGCTCGGTTAT 6246
616 ..TyrThrAlaValValTyrHisAsnArg.....Se 625
   ||| ||| ::| |||
6247 TATTATGAAGTGAATTTAGCCCATATCCGTTATGACGGCACCTGGAATAC 6296
625 rPro.....AspPheHisGluGluLys.....V 634
   ||||| ||| ::::: |||:::
6297 ACCAGTCACCTTTGATGTAACCAAAAAATATCCGATTTAATCTGGGAA 6346
634 alLysLeuPro.....AlaThrLeuThrAspHisHisLysLeu 646
   ||| ||| ::::: |||::: |||
6347 ATAAACACACCTGACTTTACTGCTCAAGCTTTCAAGCAGAGATGAATTG 6396
647 LeuPheThrPheTyrHisValSerCysGlnGlnLysGlnAsnThr.... 661
   ||| ||||| ::| ||| ::||| |||
6397 CTGGTGATGTTTATAAAAAACAAGATCAATTAAATCAATACACAACAC 6446
662 ...ProLeuGlu..... 664
   |||::: |||
6447 AGTACCAATAAAGACATATATATCACTTCCAATATGTCTTAAAGAAA 6496
665 ..ThrProVal..... 667
   |||||
6497 TGACACCTGAATAATCACAACCTAACGCTTATAACAGTTGATTAATAAT 6546
667 ..... 667
6547 AGTATTATTGGTGCAATAATCGCTAAGCAGAAAGCTACGAAATCCCTTC 6596
668 .....GlyTyrThrTrp..... 671
6597 ATCAGTAATAGTAATAACGGTTATGATGGGAGATGCGTATCTGAGTA 6646
671 ..... 671
6647 TGGTGTATGGCGAATAATATTCAGCCATCAAACTGAGTCCCTCATCAGAT 6696
672 .....IleProMetLeuGlnAsnGly..... 678
   ::| ::::: |||||
6697 AAGTTAAACTCTCACCAAGGTTAAGATTATTCATATGAGACTTTGTAGG 6746
678 ..... 678
6747 CCGACAACGCCAACCAATGCAACCTGATGAAGAATAACGGTCAGCTTGGTG 6796
679 .....ArgLeuLys 681
6797 ATAAATTATATTTATTTACTACTCTAGGTATTAACCCCAATAATTGTGCG 6846
682 ThrGlnPheCysLeuProVal..... 689
   ::| ::||| |||||
6847 AATAAAAAATTCATTTACCCCTGTTTATCAGTATAGTGGAACACTACCAA 6896
689 ..... 689
6897 TAATGAGAAAGACGCTGCTGTTTATCGAGAAAGTACTACTAATTG 6946
690 .....SerLeuGluLysProProGlnAlaTyrSerVal 700
   ::||| ::| |||:::
```



```

1183 .....GlnHisTyrLeuAlaGlyLeuValLeuThrGluL 1194
1802 ATTAACACCGCTGTGC AAAATCAGCGCGCAGAACTGATATTGACCAATC 8551
1194 euAlaVal.....IleLeuAspProAspAlaGlyLeuPhe 1206
1852 TCACATACAGGACAAACCATCCAGAGCTGATGCTGAA..... 8592
1207 GlyLeuHisLysLysValIleAsnMetValHis..AsnLeuLeuSerSe 1222
8593 .....AAAACTGTGCTAGAAAAAAACCGCGCGGAACCCAGCTCCGG 8633
1222 rHisAspSerAspProArgTyrSerAspProGlnIleLysAla..... 1236
8634 TTTTGATAGCTACAGCAATCTCAGCATGAAGACATCAACCGCGGTGA 8683
1237 .....ArgValAlaMet 1240
8684 AACAGGCAATGGCGTTCGCTCCGCTCGCTGGCATCTCTACAGCCCTT 8733
1241 LeuTyrLeuProLeuIleGlyIleMetGluThrValProGlnLeuTyr 1257
8734 CAAGCATCACATCTGGCGGCGCAGCACTTGATCTGGCGCCCAACATCTT 8783
1257 rAspPheThrGluThrHisAsnGlnArgGlyArgProIleCysIleAlaT 1274
8784 CGGCTTCGCTGATGTGGCAGCCGTGGGGG..GCGATCGCCCAAGCCA 8830
1274 hrAspAspTyrGluSerGluSerGlySerMetIleSerGlnThrValAla 1290
8831 CAGGTAAATGTCATGAGTCTCCGCGCAGTGTATGAA..... 8868
1291 MetAlaIleAlaGlyThrSerValProGlnLeuThrArgProGlySerPh 1307
8869 .....ACCGAAGCGGATAAAATCAGCCAATCTGAAGCCTA 8903
1307 eleuLeuThrSerThrSerGlyArgGlnHisThrThrPheSerAlaGluS 1324
8904 C.....CGTCGGCGTCTCAGGAA..... 8922
1324 erSerArgSerLeuLeuIleCysLeuLeuTyr..ValLeuLysAsnAla 1339
8923 .....TGGAAATTCAGCGTAAATAC 8943
1340 AspGluThrValLeuGlnLysTrpPheThrAspLeuSerValLeuGlnLe 1356
8944 GCCGAGGCAGAGCTGAAACAAATCGATGCTCAACTTGTGCTGGCGCAGT 8993
1356 uAsnArgLeuLeuAspLeuLeuTyrLeuCysValSerCysPheGluTyrL 1373
8994 GCGCGGTGAAGCCGCAATATG.....CAGAAAA 9022
1373 ysgGlyLysLysValPheGluArgMetAsn....SerLeuThrPheLys 1387
9023 CCAGCCTAAAAACCCACAAAGAGCAGACTCATGCACAACCTGACTTCCTG 9072
1388 LysSerLysAspMetArgAlaLysLeuGluGluAlaIleLeuGlySerI 1404
9073 CAACATTAAGTTCAGTAATCAGGCGCTGTACAACACTGGCTGCTGTCGATT 9122
1404 eGlyAla.....A 1407
9123 GTCCGCCATTCTACTTCAGTTCTATGATTTAACGGTAGCTCGCTTTGA 9172
1407 rgGlnGluMetValArgArg.....SerArg... 1415
9173 TGGCGGAAATGGCCTATCGCTGGAGACTAACGATACCGCATCACGCTTT 9222
1416 .....GlyGlnLeuGluArgSerProSerGlySerAlaPheGlySe 1429
9223 ATCAAAACCCGCGCGCTGGCAGGGAACCCATGCCGCTTTCCTCCGCGGTGA 9272
1429 r.....GlnGluAsnLeuArgTTPArgLysAspMetThrHisTTPArgG 1444

```

9273	AACTTAATGCTGAATCTGGCACAGATGGAAGAT..GCCCACTGAAAC	9319
1444	InasnThrGluLysLeuAspLysSerArg.....AlaGluIle	1456
9320	AGGATAAACGCGTACTGGAGGTAGAACGTACC GTTCGTGGCGCAAGTC	9369
1457GluHisGluAl	1460
9370	TATGCCAAATTTACCGCAAGATAAATTTATCCTGACTCAGCAATATAGAGAA	9419
1460	aleuIle.....AspGlyAsnLeuAlaThrGluAlaasnleuIleI	1474
9420	GTTGGTGAGTAAGGTTCAAGCGCAGCGCCGCAAGACAATAAATAGCTGG	9469
1474	IleLeuAspThrLeuGluIleValGlnThrValSerValThrGluSer	1490
9470	CGTTTAGTACCAATATCCATCTCTAGAACGCTCCATTTCGTTATCT	9519
1491	LysGluSerIleLeuGlyGly.ValLeuLysValLeuLeuHisSerMet	1507
9520	ACCTTGAACATTAGCAGCGGATATCTGATTCATTGGTAAACCCGTCG	9569
1507	IacLysAsnGlnSerAlaValTyrLeuGlnHisCys	1518
9570	TATTTAAACAGATCAGCGTTACCGCTGCCAGCACTGC	9604

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-952-127-11

seq_documentation_block:

; Sequence 11, Application US/08952127

; Patent No. 6211336

GENERAL INFORMATION:

APPLICANT: Shiloh, Yosef

APPLICANT: Tagle, Danilo A.

APPLICANT: Collins, Francis S.

TITLE OF INVENTION:	ATAXIA-TELANGIECTASIA GENE
NUMBER OF SENTENCES:	34

NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:CORRESPONDENCE ADDRESS:
ADDRESSEE: Kohn & Associates

ADDRESS: Kohn & ASSO
STREET. 30500 NO 6211

CITY: Farmington
STREET: 30500 N
:;

STATE: Michigan

STATE: MICHIGAN
COUNTRY: U.S.A.

ZIP: 48334

211, 40004
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Rel

```

; CURRENT APPLICATION DATA:

```

APPLICATION

FILING DATE: ;

CLASSIFICATION: 530 ;

ATTORNEY/AGENT INFORMATION: ;

NAME: Kohn, Kenneth I.

REGISTRATION NUMBER: 30,995

REFERENCE/DOCKET NUMBER: 2290.00029

TELECOMMUNICATION INFORMATION: 810 630 6050
; EDITED BY: 810 630 6050

TELEPHONE: 810-539-5050
FAX: 810-530-6000

TELEFAX: 810-539-5055
INFORMATION FOR SEQ ID NO: 11.

```

; INFORMATION FOR SEQ ID NO: 11:
: SEQUENCE CHARACTERISTICS:

```

SEQUENCE CHARACTERISTICS:

LENGTH: 9620 base pairs
TYPE: nucleic acid

TYPE: nucleic acid
STRANDEDNESS: single

```

; STRANDEDNESS: SINGLE
; TOPOLOGY: linear

```

TOPOLOGY: Linear
MOLECULE TYPE: cDNA

MOLECULE LIFE: CDNA
ORIGINAL SOURCE:

ORGANISM: Mus musculus

POSITION IN GENOME:
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

CHROMOSOME/SEGMENT: Chromosome 9, Band 9C

US-08-952-127-11

US-08-952-127-11


```

1479 uileValVal..... 1482
      ::: ::
6254 GCTATGTTAGAAAAAGACAATGAATAAAGCAGAAGGTATTGTCTGCC 6303
1483 ....GlnThrValSerValThrgluSerLysgluSertleuglygly 1497
      ::: |||||::: ||::: || |
6304 TTGAAGAAGAACTCTCACTGGTCCAAAGTAGAGAAAACAGCTTCGTGA 6353
1498 ValLeuLysValLeu..... 1502
      || ::::
6354 GAATTAGATACTATGTCAAATAAACACAGGCACGTGCATCACTTCTCTGA 6403
1503 .....LeuHisserMetAlaCysA 1509
      :::: ||::: |||
6404 AAAAATGAAGGAGAAACACACAGACCTTGAGTCTCATCAAAAGTAGTGT. 6452
1509 snGInSerAlaValTyrlEuGlnHisCysPheAlathrGlnArgAlaleu 1525
      :: |||||::: :: |||
6453 .....CTCCATTGCATTCCAGGTGGCAGAGGCACAGAG 6482
1526 ValSerLysPheProGluLeuLeuPheGluGluGluThrgluGlnCysAl 1542
      || ::||:::|||
6483 GTG.....AAGAAAAGACGGAA..... 6500
1542 aaSPLeuCysLeuArgleuLeuArgHisCysSerSerSertleuGlyThrI 1559
6501 .....CTCCTTCAGACTTTGTCTCCTCGATGTCAGTGAGC 6534
1559 le.....ArgSerHisProSerAlaSerLeuTyrlLeuMetArg 1572
6535 TGTTTAAAGACAAACTCATCTCCACAGAAAGCTGCAGAGTTTGGAAAG 6584
1573 GluAsn.....PheGluileGlyAsnAsnPheAl 1582
      ::::: ||::: |||::: ||
6585 GACTCACAGGCACGTGCTTTGACAAATGTGAGCTGGAACCAATTTGC 6634
1582 aArgVal..... 1584
      |::::
6635 ACAACTGAATTAAGAGAAAGAAATTGCTTGTCAAGGAATCTGAAGCCTGC 6684
1585 .....LysMetGlnValPrometSer 1591
6685 AGGCCAGACTGAGTGAATCAGATTATGAAAAGCTGAATGTCTCCAAGGCC 6734
1592 LeuSerSer.....LeuValGl 1597
      ||::::: ||::: ||
6735 TTGCAGGCCCGCACTGCTGAGAAAGGTGAGTTCGCAATTGAGGCTGAGCTC 6784
1597 yThrSerGlnAsnPheAsnGluIuPheLeuArgArgSerLeuLysThri 1614
      :|||:::~::~: |||||~::~:~:
6785 AACACAGAGAGGAAGTGCATCAG.....CTGAGAAGAGGCATCGAGAACC 6828
1614 lEleuthrTyrrAlaglGluAsp..... 1621
      :: :::: ||| |||
6829 TGAAGAGTTCCGCAATTGAGGCCCGATGAATAAGAACGACGTCGCACATCCACAGAG 6878
1622 ..leuGluLeuArgGlu.....ThrThrpheProaspGlnValGl 1634
      ||::: ~|||~| ::::: |||~|||~:
6879 AAACCTGAAGAAGACGCGAGCGGAGAAATGATTCTTAAGGATTAAGTTGA 6928
1634 naspleuValPheasnLeuHisMet.....Ilel 1644
      ::|||~| ::|||~|||
6929 GAACCTTGAAGAAGGAATTGCAGATGTCAAGAAGAAAACGAGAGCTAGTGA 6978
1644 euSerAsprThrValLysMetLysGluHisGlnGluAspProGluMetLeu 1660
      :: |||~: :: || ~: ||
6979 TTCCTGATGCGCAGAGATTCCAAAGCAGAGAGTAGAGACTCTAAACACACAA 7028
1661 lleaspleuMetTyrrArgileAlalalysGlyTyrgln..... 1672
      |||~: ||| ||| ||| :::::
7029 ATGAGAAGAGATGGCCAGAAAGCCTGAAAGATTTTTGAATTAGACCTTGTCAAC 7078
1673 .....ThrserProGluArgLeuThr..TripleuGlnasnMetalag 1686
```

7079	GTTAAGCTCTGAAAAAGAAATCTGACAAACAATACAGAAGAAAAACAAG	17022
1686	lylyshissergluargserasnhisalaglualalaglincysleuval	17022
7129	GTCAGTTGTCAGAACTAGACAAAG.....	7151
1703	hisseralaalaleuvalalaglutyrlleusermettleugluaspargly	1719
7152TTACTCTCTTCATTTAAGAAGTCTGTAGAGAAGAAAGA	7189
1719	styrleuprovalglycysvalthrpheglhnsnilesersersaval	1736
7190	GCAAGCAGAGATA.....CAGATCA	7209
1736	eugluagluseralavalseraspvalvalserproaspgluglugly	1752
7210	AAGAGAATCTAAAACTGCAGTGCAGATGCTT.....	7241
1753	ilecysserglylystyrphethrgluserglyleuvalglyleuengl	1769
7242CAGATCAGTTAAAGGAGCTAAATGA	7267
1769	uglnalaalalaserphesermetaglmettyrgluavalalasng	1786
7268	GCGCAGTACGAGCC.....TTGTGTGTGACCAAGAAATTATGAAGG	7308
1786	luvaltyrlylsvalleu..ileprolehisglualaasnargaspala	1801
7309	CCACGAAACAGAGCTTAGACCCACCAATAGAGGAAGAGCATCAGCTGAGA	7358
1802	lyslylseuserthrilehisglylsleuengl	1812
7359	AATGACATTGAAAAAGCTGAGAGCCCGCTAGAA	7391

seq_name: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq: PCT-US95-16216-2

```

seq_documentation_block:
; Sequence 2, Application PC/TUS9516216
; GENERAL INFORMATION:
; APPLICANT: Yen, Timothy J.
; APPLICANT: Ratner, Jerome B.
; TITLE OF INVENTION: Nucleic Acid Encoding a Transiently
; TITLE OF INVENTION: Expressed Kinetochores Protein, and Methods of Use
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dann, Dorfman, Herrell and Skillman
; STREET: 1601 Market Street Suite 720
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103-2307
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/16216
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/353,700
; FILING DATE: 09-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Reed, Janet E.
; REGISTRATION NUMBER: 36,252
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 563-4100
; TELEFAX: (215) 563-4044
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10136 base pairs

```

```

;      TYPE: nucleic acid
;      STRANDEDNESS: double
;      TOPOLOGY: linear
;      MOLECULE TYPE: DNA (genomic)
;      HYPOTHETICAL: NO
;      ANTI-SENSE: NO
PCT-US95-16216-2
```

```

alignment_scores:
      Quality: 158.50      Length: 2111
      Ratio: 0.171      Gaps: 93
Percent Similarity: 44.008      Percent Identity: 17.859
```

alignment_block:
US-09-737-246-2 x PCT-US95-16216-2 ..

Align seg 1/1 to: PCT-US95-16216-2 from: 1 to: 10136

```

72 AspSerGlyProLeuArgAspLeuIleGluPheProAspAspIleG1 88
   ::::| | | | | | | | | | | | | | | | | | | | | |
1950 GAGTCCAAAGCCTTGCTGAGTGCCTTAGAGTTAAAAAGAAAGATATGA 1999
      88 uValValTyrSerProArgAspCysArgThrLeuValSerAlaValProG 105
      | :::: | | | | | | | | | | | | | | | | | | | |
2000 AGAATTG.....AAAGAAGAGAAAACCTGTTTCTTGTGGAAAA 2040
      105 IuGluSerGlu.....MetAspPro..... 111
      ::| | | | | | | | | | | | | | | | | | | |
2041 GTGAAAAACGAAAAACITTTTAACCTCAGATGGAATCAGAAAAAGAAACTTG 2090
      112 .....HisValArgAspCysIleArgSerTyr..... 120
      | | | | | | | | | | | | | | | | | | | |
2091 CAGAGTAAAAATTAACTACTTGGAACCTGCTGGAAGACACAGCAAAATAAA 2140
      121 .....ThrGluA 123
      | | | | | | | | | | | | | | | | | | | |
2141 AAGTCATGAATACAACGAGAGAGTAAGAACGCTGAGATGACAGAGAAAA 2190
      123 sPTPrAlaIleValIleArgLysTyrHisLysLeuGlyThrGlyPheAsn 139
      :: :::: | | | | | | | | | | | | | | | | | |
2191 ACCTAAGTGTGAGATCAGAAACCTTCAC..... 2219
      140 ProAsnThrLeuAspLysGlnLys.....GluArgGlnLysGlyLe 153
      | | | | | | | | | | | | | | | | | | | |
2220 ...AACGTGTTAGACAGTAAGTCAGTGAGAGGTAGAGACCAGAAACTAGC 2266
      153 uProLysGlnValPheGluSerAspGluAlaProAspGlyAsnSerTyrG 170
      :::: | | | | | | | | | | | | | | | | | |
2267 TTATATGGAGCTACAGAGAAAAGCTGAGTTCTCAGATCAGAAAACATCAGA 2316
      170 IAspAspGln..... 173
      :::: | | | | | | | | | | | | | | | | | |
2317 AGGAAATAGAAAAATATGTGTTGAAGACTTTCAGCTTACTGGGCAAGTT 2366
      174 AspAspLeuLysArgArgSerMetSerIleAspAspThrProArgGlySe 190
      ::| | | | | | | | | | | | | | | | | | | |
2367 GAAGATCTAGAACACAGCTTCAGTTACTGTCAAATGAATAATGACAAA 2416
      190 rTrpAlaCys.....SerIlePheAspL 198
      : | | | | | | | | | | | | | | | | | |
2417 AGACCGGTGTTACCAAGACTTGCATGCCGAATATGAGAGCCTCAGGGATC 2466
      198 euLysAsnSerLeuProAspAlaLeuLeuProAsnLeuLeuAspArgThr 214
      | | :::: | | | | | | | | | | | | | | | |
2467 TGCTAAAAATCCAAAGATGCTTCTCTGGTGACAAAT..... 2501
      215 ProAsnGluGluIleAspArgGlnAsnAspAspGlnArgLysSerAsnAr 231
      :::: | | | | | | | | | | | | | | | |
2502 .....GAAGATCATCAGAGAAAGTCTTTTGGC 2527
      231 gHisLysGluLeuPheAlaLeuHisProSer..... 241
      :::: | | | | | | | | | | | | | | | |
```

```

2528 TTTTGATCAGCAGCCCTGCCATGCATCATTCCTTTGCAAAATATTAATTGGAG 2577
      242 .....ProAspGluGluGluProIleGluArgLeuSerVal 253
      | | | | | | | | | | | | | | | | | | | |
2578 AACAAAGGAGCATGCCTTCAGAGAGAGTGAATGT...CGTTTAGAAGCA 2624
      254 ProAspIleProLysGluHisPheGlyGlnArgLeuLeuValLysCysLe 270
      :::: | | | | | | | | | | | | | | | |
2625 GACCAAGTCCGAAAAAT.....TCTGCCATCCTTCAAAATAGAGTTGA 2668
      270 uSerLeuLysPheGluIleGluIleGluProIlePheAlaSerLeuAla 287
      | | | | | | | | | | | | | | | | | | | |
2669 TTCACCTGAATTTTCATTAGAGTCTCAAAAACAGATGAACTCAGACCCTGC 2718
      287 euTyrAspValLysGlu.....LysLysLysIleSerGluAsn 299
      :::: | | | | | | | | | | | | | | | |
2719 AAAAGCAGTGTGAAGAGTTGGTCCAATCAAAAGAGAAATAGAAAGAAAT 2768
      300 PheTyrPheAspLeuAsnSerGluGlnMetLysGlyLeuLeuArgProHi 316
      :::: | | | | | | | | | | | | | | | |
2769 CTC.....ATGAAGCAGAACAGATG.....CA 2791
      316 sValProProAlaAlaIleThrThrLeuAlaArgSerAlaIlePheSerI 333
      | :::: | | | | | | | | | | | | | | | |
2792 TCAAAAGTTTGTGGCTGAAACAGTCAAGCGCATTAGTAAGTTACAGGAAG 2841
      333 IeThrTyrProSerGlnAspValPheLeuValIleLysLeuGluLysVal 349
      | | | | | | | | | | | | | | | | | |
2842 ACACCTTCTGCTCACCAGAAATGT..... 2864
      350 LeuGlnGlnLysAspIleGlyLysAlaGluProTyrMetIlePheLe 366
      2865 .....GTTGCTGA 2872
      | | | | | | | | | | | | | | | | | | | |
366 sGluAlaAspAlaThrLysAsnLysGluLysLeuGluLysLeu..... 380
      : :::: | | | | | | | | | | | | | | | |
2873 AACCTTAAGTGCCTTGAGAACAGAAAAAGAGCTGCAACTTTTAAATG 2922
      381 ..LysSerGlnAlaAspGlnPhe...CysGlnArgLeuGlyLysTyrArg 395
      | | :::: | | | | | | | | | | | | | | | |
2923 ATAAAGTAGAAAACCTGAGCAGCAGAGATTCAAGAATTAATAAAG..... 2966
      396 MetProPheAlaTrpThrAlaIleHisLeu..... 405
      2967 .....AGCAACCATCTACTTGAAGACTCTCTAAAGGA 2998
      406 .MetAsnIleValSerSerAlaGlySerLeuGluArgAspSerThrGluV 422
      :::: | | | | | | | | | | | | | | | |
2999 GCTACAACTTTTATCCGAACCCCTTAAGCTTGAGAGAGAAAGAAATGAGTT 3048
      422 alGluIleSerThrGlyGluArgLysGlySerTrpSerGluArgArgAsn 438
      | | | | | | | | | | | | | | | | | |
3049 CCATCATTTTCTTTAAATAAAGGAAATTGAAAGAGCTGACCCAGAGAGAA 3098
      439 SerSerIle.....ValGlyArgArgSerLeuGluAr 449
      :::: | | | | | | | | | | | | | | | |
3099 GGGACTCTTAAAGAAATTAATGCATCCTTAAATCAAGAGAAAGATGAAC 3148
      449 gThrThrSerGlyAspAspAlaCysAsnLeuThrSerPheArgProAla 466
      :::: | | | | | | | | | | | | | | | |
3149 AATCCAGAAAAAGTGAGAGTTTGCAAACTATATAGATGAAGGAGAGAAA 3198
      466 hrLeuThr...ValThrAsnPhePheLysGlnGluGly..... 477
      :::: | | | | | | | | | | | | | | | |
3199 GCATTTCAAGAGTTATCTGATCAGTACAAAGCAAGAAAACTTATTTACTA 3248
      478 AspArgLeuSerAsp.....GluAspLeu...TyrLysPh 488
      :::: | | | | | | | | | | | | | | | |
3249 CAAAGATGTGAAGAAACCGGAAATGCATATGAGGATCTTAGTCAAAAATA 3298
      488 eleuAlaAspMetArgArgProSerSerValLeuArgArgLeuArgProI 505
      : | | | | | | | | | | | | | | | |
3299 CAAAGCAGCACAGAAAGAAATTTCTAAATTA..... 3329
```

505 leThrAlaGlnLeuLysIleAspIleSerProAlaProGluAsnProHis 521
3329 3329
522 TyrCysLeuThrProGluLeuGlnValLysLeuTyrProAspSerAr 538
||||| ||| : : : : :
3330 GAATGCTTGCTAAATGAATGCACTAGTCTT.....TGTGAAATAG 3370
538 gValArgProThrArgGluIleLeuGluPheProAlaArgAspValTyrV 555
| : : : : : ||| ||| : : : : :
3371 GAAAAATGAGTTGGAACAGCTTAAGGAAGCATTTGCAAAAGAA..... 3413
555 alProAsnThrThrTyrArgAsnLeuLeuTyrIleTyrProGlnSerLeu 571
: : : : : : ||| : : : : :
3414CACCAAGATCTCTA.....ACAAATTA 3437
572 AsnPheAlaAsnArgGlnGlySerAlaArgAsnIleThrValLysVal.. 587
||||| :
3438 GCATTGCTGCAAGAAAGAAAT.....CAGAATCTGATGCTAGAGTTGGA 3481
588GlnPheMetTyrGlyGluAspProSerAsnAlaMetProV 601
||| : : : : : ||| : : : : :
3482 GACAGTCGACGCAAGCTCTGAGATCTGAGATGACAGATAACCAAAACAAT. 3530
601 alIlePheGlyLysSerSerCysSerGluPheSerLysGluAlaTyrThr 617
: : : ||||| : : : : : : : : : : : : : : : : |||
3531TCTAAGAGCGAGGCTGCTGTTTAAAGCAGAAATCATGACT 3572
618 AlValValTyrHisAsnArgSerPro.....AspPheHisG1 630
: : : ||| : : : : : ||| : : : : :
3573 TTTAAGGAGAACAACAAATGCCAAAGGAAGTTAATGACTTATTACA 3622
630 uGluIleLysValLysLeuProAlaThrLeuThrAspHisHisLeuL 647
: ||| : : : : : ||| |||
3623 AGAGAAATGACACAGCTGATGAGGTAATGAAGACTAAACAT..... 3662
647 eupheThrPheTyrHisValSerCysGlnGlnLysGlnAsnThrProLeu 663
: : : ||||| : : : : : : : : : : : : : : : : ||| : : :
3663GAATGTCAAAATCTAGAAATCAGAAACCAATT 3692
664 GluThrProValGlyTyrThrTrpIlePrometLeuGlnAsnGlyArgLe 680
: : : : : ||| : : : : : ||| : : : : :
3693 AGGAACCTCTGTG.....AAAGAAAGAGA 3715
680 uLysThrGlnPheCysLeuProValSerLeuGluLysProProGlnA 697
: : : : : |||
3716 GAGTGAGAGAAATCAATGT..... 3734
697 laTyrSerValLeuSerProGluValProLeuProGlyMetLysTrpVal 713
3734 3734
714 AspAsnHisLysGlyValPheAsnValGluValAlaValSerSerIl 730
||| ||| : : : : : ||||| : : : |||
3735 ..AATTTAAACCTCAGATGCATCTTGAAGTTAAAGAAATTTCT.... 3776
730 eHisThrGlnAspProTyrLeuAspLysPhePheAlaLeuValAsnAlaL 747
||||| : : : : : ||| : : : : :
3777CTAGATAGTTATATATGCGCAGTTGGTGCAAT 3807
747 euAspGluHisLeuPheProValArgIleGlyAspMetArgIleMetGlu 763
||| : : : ||| : : : : : : : : : : : : : : : : |||
3808 TAGAAGCTATGCTA.....AGAAATAAGGAATTAAACTTCAGGAA 3848
764 AsnAsn.....LeuGluAsnGluLeuLysSerSerIleSe 775
: : : : : ||| : : : : ||||| : : : : : : : : : : :
3849 AGTGAGAGAGAGAGAGAGTGCCCTGCAGCATGAATTCACAGACAATTAGAGG 3898
775 rAlaLeuAsnSerSerGlnLeuGluProValValArgPheLeuHisLeuL 792
: ||| : : : : : ||| : : : : : : : : : : : : : : : :
3899 AGATCTTGAAACCAGCAATTTGCAAGACATGCAG..... 3932

792 euLeuAspLysLeuIleLeuLeuValIleArgProProValIleAlaGly 808
3933TCACAA 3938
809 GlnIleValAsnLeuGlyGlnAlaSerPheGluAlaMetAlaSerIleIl 825
: : : ||| : : ||| : : : : : : : : : : : : : : : : |||
3939 GAATTTAGTGCCCTTAAGACTGTGAATATAGATGCCGAGAGAAAGTATAT 3988
825 eAsnArgLeuHisLysAsnLeuGluGlyAsnHisAspGlnHisGlyArgA 842
| : : : : ||| : : : : : : : : : : : : : : : :
3989 TTCAGGGCCCTCATGAGTTGTCAACACACTCAAAACGACAAATGCACACCTTC 4038
842 snSerLeuLeuAlaSerTyrIleHisTyrValPheArgLeuProAsn... 857
: : : ||| : : : : : : : : : : : : : : : :
4039 AGTGCTCTCTGCCAAACAACAATGAACAAGCTGAATGAGCTTAGAGAAATA 4088
858ThrTyrProAsnSe 862
4089 TGTGAATACTGCAGGCTGMAAAGTATGAACCTCGTAACCTGACCTGAATGA 4138
862 rSerSerProGlyProGlyGlyLeuGlyGlySerValHisTyrAlaThrM 879
: ||| : : : : : ||| : : : : :
4139 TTCAAGGTCA.....GAATGTATCACAG 4161
879 etAlaArgSerAlaValArgProAlaSerLeuAsnLeuAsnArgSerArg 895
: : : ||| : : : : : : : : : : ||||| : : : : :
4162 CAACTAGCAAAATGGCAGAGAGGTAGGAAACTACTAAATGAGTTAAA 4211
896 SerLeuSerAsnSerAsn.....ProAspIl 904
||||| : : : : : : : : : : :
4212 ATATTAAATGATGACAGTGGTCTTCTCCATGCTGAGTTAGTGAAGACAT 4261
904 eSerGlyThrProThrSerProAspAspGluValArgSerIleIleGlys 921
| : ||| : : : : : : : : : : ||||| : : : : :
4262 ACCAGAGGTGAATTTGGTGAAACAACCAATGAACAGACACCCTGTGCTT 4311
921 erLysGlyLeuAspArgSerAsnSerTrpValAsn..ThrGlyGlyPro 936
||||| : : : ||||| : : : : : : : : : : ||| : : :
4312 TGGCTCCATTTGGACGAGAGTAATTCCTACGAGCAGCTTGACATTTGCAGAC 4361
937 Lys.....AlaAlaProTrpGlySerAsnProSerPr 947
||| : : : : : ||| : : : : :
4362 AAAGAGTTCAAAATGCACCTTGGCCGAATTCGAAGAGAAATTTATCTTT 4411
947 oSerAlaGluSerThrGlnAlaMetAspArgSerCysAsnArgMetSers 964
: : : : : ||| : : : : : ||| : : : : :
4412 ACAAAAGTGAACACAAATTTTACATGATCAGCAGCTGTCAGATGAGCTCTA 4461
964 erHisThrGluThrSerSerPheLeuGlnThrLeuThrGlyArg..... 978
: : : : : ||| : : : : : : : : : : : : : : : :
4462 AAATGTCAGAGCTGCAGACCTATGTTGACTCATTTAAAGGCCGAAAAATTTG 4511
979LeuProThrLy 982
4512 GTCTTGTCAACGAATCTGAGAAACTTTCAAGGTGACTTGGTGAGAGAT 4561
982 sLysLeuPheHisGluGlu.....LeuAlaLeuGlnTrpV 994
: : : ||| ||||| : : : : : ||| : : : : :
4562 GCAGCTGGGCTTGGAGGAGGGGCTGTTCCATCCCTGTCATCCCTTGTG 4611
994 alValCysSerGlySerValArgGluSerAlaLeuGlnGlnAlaTrpPhe 1010
||| : : : ||| : : : : : ||| : : : : : |||
4612 TGCCTGACAGCTCTAGTCTT....AGCAGTTTGGAGACCTCCCTT 4655
1011 phePheGluLeuMetValLysSerMetValHisHisLeuTyrPheAsn.. 1026
: : : ||| : : : : : : : : : : ||| : : : : :
4656 TACAGAGCTCTTTTGAACACAGACAGAGATATGCTCTTTTGGAGTAATTT 4705
1027 .AspLysLeuGluAlaProArgLysSerArgPheProGluArgPheMetA 1043
: : : : : : : : : : : ||| ||| : : : : :
4706 AGAAGGGGCTGTTTACGCAAAACAGTGAGTGATGAAGTATTTTGA 4755
1043 spAspIleAlaIleValSerThrIle...AlaSerAspIleValSer 1058

4756 GCAGTCTGCAGACCTAATGTTGACTCATTAAGCGCAAAATTTGGTC... 4802
1059 ArgPheGlnLysAspThrGluMetValGluArgLeuAsnThrSerLeuAl 1075
4803TTGTCACGAATCTGAG 4819
1075 apPhePheLeuAsnAspLeuSerValMet.....AspArgG 1088
4820 AAACCTTCAAGGTGACTTGGTGAAGAGATGCAGCTGGCCTTGAGAGAG 4869
1088 lypPheValPheSerLeuIleLysSerCysTyrLysGlnValSerSerLys 1104
4870 GGCTCGTCCATCCCTGTCATCCTCTTGTTGCTGAC...AGCTCTAGT 4916
1105 LeuTyrSerLeuProAsnProSerValLeuValSerLeu.....ArgLe 1119
4917 CTTCAGCAGTTTGGAGACCTCCTTTTACAGAGCTCTTTTGAACAGAC 4966
1119 uasPheLeuArgIleIleCysSerHisGluHisTyrValThrLeuAsn. 1135
4967 AGGAGATATGCTCTCTTTGAGTAATTAGAAGGGGTGTTTCAGCAAAACC 5016
1136LeuProCysSerLeu..... 1141
5017 AGTGCAGTGTAGATGAAGTATTTTGCAGCAGCTCTGCAGAGAGAATCTG 5066
1142ThrProAlaSerProSerProSerValSerSerAlaTh 1155
5067 ACCAGGAAAGAAACCCCTTCGGCCCCAGCGAAGGGTGTGAAGAGCTTGA 5116
1155 rSerGlnSerSerGlyPheSerThrAsnValGlnAspGlnLysIleAla 1172
5117 GTCCCTCTGTGAGGTGTACCGCAGTCCCTCGAG..... 5150
1172 smetPheGlnLeuSerValProPheArgGlnGlnHisTyrLeuAlaGly 1188
5151AAGCTAGAGAGAAATGGAAAGTCAAGCGATTTATGAAAAAT 5192
1189 LeuValLeuThrGluLeuAlaValIleLeuAspProAspAlaGluGly 1205
5193 AAGGAAATTCAAGAGCTCGAGCAGTTATTAGTTCTGAAGCAAGAGCT 5242
1205 upPheGlyLeuHisLys.....L 1211
5243 TGACTGCTTAGGAAGCAGTATTTGTCAGAAATGAACAGTGGCAACAGA 5292
1211 ysvAlIleAsnMetValHisAsnLeuLeuSerSerHisAspSerAspPro 1227
5293 AGCTGACAAGCGTACTCTGAGATGAGTCCAAAGTTGGCGGAGAAAAAG 5342
1228 ArgTyrSerAspProGlnIleLysAlaArgValAlaMetLeuTyrLeuPr 1244
5343 AAACAGACGGAA...CAACTGTCACTTGAGCTGGAAGTAGCAGACTCCA 5389
1244 oleuIleGlyIleIleMetGluThrValProGlnLeuTyrAspPheThrG 1261
5390 GCTACAGGTCTGACTTAAGTTCGCGTCTTTCCTTGGCATCGACACAG 5439
1261 luthrHisAsnGlnArgGlyArgProIleCysIleAlaThrAspAspTyr 1277
5440 AAGATGCTATTCAAGGCCGAATGAGAGCTGTGACATATCAAAAGAACAT 5489
1278 GluSerGluSerGlySer.....MetIleSe 1286
5490 ACTTCAGAACTACAGAAAGAACCAAGCATGATGTTCAATCAGATTTTG 5539
1286 rGlnThrValAlaMetAlaIleAlaGlyThrSerValProGlnLeuThra 1303
5540 TGATAAAGATGCTCAGCAGACCTCAATCTAGACATTGAGAAATAACTG 5589
1303 rgProGlySerPheLeuLeuThr...SerThrSerGlyArgGlnHisThr 1318
:: |||::: ||| ::: |||||:::||||

5590 AGACTGGTGCAGTGAAAACCCACAGAGAGTGTCTGGGGAACAGTCCCCA 5639
1319 ThrPheSerAlaGlu.....SerSerArgSerLeuLeuI 1330
5640 GATACCAATTATGAGCCTCCAGGGGAGAGATAAAACCCAGGGCTCTTCAGA 5689
1330 eCysLeuLeuTrpValLeuLysAsnAlaAspGluThrValLeuGlnLysT 1347
5690 ATGCATTTCTGAATTTGTCATTTTCTGCTCTAATAGCTTTGGTACCTATGG 5739
1347 rpPhe.....ThrAspLeuSerValLeuGlnLeu..... 1356
5740 ATTTCCCTGGGAATCAGAGATATCCATTAATCTTCAACTGGCGGTAAA 5789
1357AsnArgLeuLeuAspLeuLeuTyrLeuCysValSerCysPh 1370
5790 GAGACATCAATGAGAATTGAGATTACTTCATGTG.....AT 5827
1370 eGluTyrLysGlyLysLysValPheGluArgMetAsnSerLeuThrPhe 1387
5828 AGAGACCGGTGACAGAAAAAGTTGAAGTTTGTCTAAATGAATGAAGAAT 5877
1387 ysLysSerLys.....AspMetArgAlaLysLeuGlu 1397
5878 TAGACTCAAACTCCATTTTACAGAGAGTACAACATAATGACCAAAATTGAA 5927
1398GluAlaIleLeuGlySerIleGlyAla..... 1406
5928 GCATGCATAGAAATTGGAAAAAATAGTGGGAACTTAAGAAAGAAACTC 5977
1407ArgGlnGluMetV 1411
5978 AGATTTAAGTGAATAATTGGAATATTTTCTTGATGCACACGAGGTAC 6027
1411 alArgArg.....SerArgGly.....GlnLeuGluArgSer 1421
6028 TCCAGAGAGTAGAAACTTCTGAAGGCCCTCAATCTGATTTAGAAATGCAT 6077
1422 ProSerGlySerAlaPheGlySerGlnGluAsnLeu..... 1433
6078 GCAGATAAATCA.....TCACGTGAAGATATTGGAGATTAATGTGGC 6118
1434ArgTrpArgLysAspMetThrHisTrpArgGlnAsnT 1446
6119 CAAGTGAATGACAGCTGGAAGAGAGATTTCTTGATGTGCAAAATGAGC 6168
1446 hrGluLysLeuAspLysSerArgAlaGluIleGluHisGluAlaLeuIle 1462
6169 TGAGTAGGATCAGATCGGAGAAAGCTAGCATTGAGCATGAAGCCCTCTAC 6218
1463 AspGlyAsnLeuAlaThrGluAlaAsnLeuIleIleLeuAspThrLeuGl 1479
6219CTGAGGCTGACTTAGAGGTAGTTCAAACAGAGAA 6253
1479 uIleValVal..... 1482
6254 GCTATGTTTAGAAAAAGACAATGAAAAATAGCAGAGGTTATGTCTGCC 6303
1483GlnThrValSerValThrGluSerLysGluSerIleLeuGlyGly 1497
6304 TTGAAGAAGAACPTCAGTGTGCACAAAGTGAGAGAAACCAGCTTCGTGGA 6353
1498 ValLeuLysValLeu..... 1502
6354 GAATTAGATACTATGTCAAAAAAACACAGGCACTGGATCAGTTGTCGA 6403
1503LeuHisSerMetAlaCysA 1509
6404 AAAAATGAAAGAGAAACACAGAGAGCTTGAGTCTCATCAAAAGTGAAGT 6452
1509 snGlnSerAlaValTyrLeuGlnHisCysPheAlaThrGlnArgAlaLeu 1525
6453CTCCATTGCATTCAGGTGGCAGAGCGCAGAG 6482

1526 valserLysPheProGluLeuLeuPheGluGluThrGluGlnCysAl 1542
||| ::|||::|||
6483 GTG.....AAGGAAAGACGAA..... 6500
1542 aaspleuCysLeuArgLeuLeuArgHisCysSerSerIleGlyThrI 1559
|||::| ::|||::|
6501CTCCTTCAGACTTGTCTCTGATGTGAGTGAGC 6534
1559 le.....ArgSerHisProSerAlaSerLeuTyrLeuLeuMetArg 1572
::| ::|||::|
6535 TGTAAAGACAAACTCATCTCCAGGAAAGCTGCAGAGTTTGGAAAG 6584
1573 GlnAsn.....PheGluIleGlyAsnAsnPheAl 1582
::|::| ::|||::|
6585 GACTCACAGGCACTGTCTTTGACAAATGTGAGCTGGAACCAATTC 6634
1582 aargVal..... 1584
|::|::|
6635 ACAACTGAATAAGAGAAAGAAATTGCTTGTCAAGGAATCTGAAGCCTGC 6684
1585LysMetGlnValPrometSer 1591
|||::|::|
6685 AGGCCAGACTGAGTGAATCAGATTATGAAAGCTGAATGTCTCCAGGCC 6734
1592 LeuSerSer.....LeuValG1 1597
|||::|::|
6735 TTGGAGGCCCGCACTGTGTGAGAAAGGTGAGTTCGCATTGAGGCTGAGCTC 6784
1597 yThrSerGlnAsnPheAsnGluGluPheLeuArgArgSerLeuYsThrI 1614
::|::|::|::|::|
6785 AACACAGAGGAGGAGTGCATCAG.....CTGAGAAAGAGGCATCGAAGAAC 6828
1614 leLeuThrTyrAlaGluGluAsp..... 1621
::|::| ::|||
6829 TGAGAGTTGCAATGATGAGCCGATGAAAGAGACAGCTGCACATCGCAGAG 6878
1622 ..leuGluLeuArgGlu.....ThrThrPheProAspGlnValG1 1634
|||::| ::|||
6879 AAACGTGAAGAACGCGGAGCGGAGAAATGATTCACCTTAAGATTAAGTTGA 6928
1634 naspleuValPheAsnLeuHisMet.....IleL 1644
::|::| ::|||::|
6929 GAACCTTGAAGGGAATTCAGATGTCCAGAGAAACACAGAGCTAGTGA 6978
1644 euserAspThrValLysMetLysGluHisGlnGluAspProGluMetLeu 1660
::|::|::|::|
6979 TTCTTGATGCCGAGAAATTCCAAAGCAGAGTAGAGACTCTAATAAACACAA 7028
1661 ileAspleuMetTyrArgIleAlaLysGlyTyrGln..... 1672
|||::| ::|||
7029 ATAGAGAGATGGCCAGAACCTGAAGATTTTGAATTAGACCTTGTCAAC 7078
1673ThrSerProGluArgLeuThr...TrpLeuGlnAsnMetAlaG 1686
::|::| ::|||::|
7079 GTTAAGGTCTGAAAAAGAAATCTGACAAAAACAATACAGAAACAAAG 7128
1686 lYlYSHisSerGluArgSerAsnHisAlaGluAlaAlaGlnCysLeuVal 1702
|||::| ::|||
7129 GTCAGTGTCTCAGAACTAGACAAG..... 7151
1703 HisSerAlaAlaLeuValAlaGluTyrLeuSerMetLeuGluAspArgLy 1719
|||::|::|::|
7152TTACTCTCTTCATTTAAAAGTCTGTAGAGAAAAAGGA 7189
1719 sTyrLeuProValGlyCysValThrPheGlnAsnIleSerSerAsnValL 1736
::|::|
7190 GCAAGCAGAGATR.....CAGATCA 7209
1736 euglUGluSerAlaValSerAspAspValValSerProAspGluGluGly 1752
|||::|::|
7210 AAGAAGAATCTAAAACTGCAGTGTGCTT..... 7241

1753 ileCysSerGlyLysTyrPheThrGluSerGlyLeuValG1LeuGlu 1769
::|::| ::|||
7242CAGAAATCAGTTAAAGAGACTAAATGA 7267
1769 uGlnAlaAlaAlaSerPheSerMetAlaGlyMetTyrGluAlaValAsnG 1786
::|::|::|
7268 GGCAGTAGCAGCC.....TTGTGTGTGACCAAGAAATTATGAGAG 7308
1786 luValTyrLysValLeu...IleProIleHisGluAlaAsnArgAspAla 1801
::|::| ::|||
7309 CCACAGAACAGAGTCTAGACCCACCAATAGAGGAGAGCATTACAGCTGAGA 7358
1802 LysLysLeuSerThrIleHisGlyLysLeuGln 1812
::|::|::|
7359 AATAGCATTTGAAAGCTGAGAGCCCGCCTAGAA 7391

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-471-119A-1

seq_documentation_block:
; Sequence 1, Application US/08471119A
; Patent No. 5827706
; GENERAL INFORMATION:
; APPLICANT: Leitner, Ernst
; APPLICANT: Schneider, Elisabeth
; APPLICANT: Schoergendorfer, Kurt
; APPLICANT: Weber, Gerhard
; TITLE OF INVENTION: Cyclosporin Synthetase
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5827706artis Corporation
; STREET: 59 Route 10
; CITY: East Hanover
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07936
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/471,119A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: kassenoft, Melvyn
; REGISTRATION NUMBER: 26,389
; REFERENCE/DOCKET NUMBER: 100-8029/CONT/CONT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 503 8474
; TELEFAX: 201 503 8807
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 46899 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Tolypocladium niveum
; STRAIN: ATCC 34921
US-08-471-119A-1

alignment_scores:
Quality: 158.50 Length: 1789
Ratio: 0.209 Gaps: 95
Percent Similarity: 42.314 Percent Identity: 19.955

alignment_block:
US-09-737-246-2 x US-08-471-119A-1 ..

||||| : : : : :
13740 CAGCTGCAATAGTGGGTGAGCAACTGTGATAGTGCCCGGCGAGCT 13789
1378 hegluargmetasnerleuthrphelysserlyaspmetargala 1394
: : : : :
13790 ACTCACGGATC.....TGGCGGCGCTTCGATCC 13818
1395 lysleugluAlaIleleuglyserlleGlyAlaArgln..... 1408
: : : : :
13819 TCTCTGGTCAGG.....CTGGGAGATGTCGGTCACGATCGAGGAGCA 13862
1409GluMetValArgArgSerArgGlyGlnLeuGluArgSerPros 1423
: : : : :
13863 CTATACAGAAGACTTGGAGGAATCTGCCGGGTCCATCGGTTACCTCTT 13912
1423 er.GlySerAlaPheGlySerGlnGluasn.....LeuArgTr 1435
|| : : : : ||
13913 CGTGGTACTGCTTGGCGCCCTACGCGCAGCCATTATCGCCTCACAGGTT 13962
1435 pArgLysAspMetThrHisTrpArgGlnAsnThrGluLysLeuAspLys. 1451
: : : : :
13963 CCGA.....AGACGCAACTATAGGGACACCAATCGCC 13994
1452 SerArgAlaGluIleGluHisGluAlaLeuIleAspGlyAsnLeuAlaTh 1468
: : : : :
13995 AATCGTAACCGACCTGAACCTTGAGCAGATATCGGCTTCTCGTCAATAC 14044
1468 rGluAlaAsnLeuIleIleLeuAspThrLeuGluIleValValGlnThrV 1485
| : : : : |
14045 GCAATGTATACGCATT.....ACCG 14064
1485 alSerValThrGluSerLysGluSerIleLeuGlyGlyValLeuLysVal 1501
| : : : : |
14065 TCACGAGAGAGACCTTTGAGTCA..... 14090
1502 leuLeuHisSerMetAlaCysAsnGlnSerAlaValTyrLeuGln.Hisc 1518
| : : : : |
14091 CTAGTGCAGCAGGTCCGCTCAACGCGACAGCTGCATTGCCCATCAGGA 14140
1518 ySPheAlaThrGlnArgAlaLeuValSerLysPheProGluLeuLeuPhe 1534
: : : : :
14141 CGTCCCGTTCGAGAAGATCGTCTCTACTCTTTGCCCGGTTCTCGAGATG 14190
1535 GluGluGluThrGluGlnCysAlaAspLeuCysLeuArgLeuLeuArgH 1551
: : : : :
14191 CATCCCGAAACCCACTTGTGCACTCA..TGTTTGGCGGTGCATTCCGACG 14237
1551 s..... 1551
14238 AAGAACCCTCGGTGAGCTGAAGCTGAAAACGCTCACAGCGAGTTGTTC 14287
1552CysSerSerSer 1555
14288 CACGAGATCACAGCCGCTTCGATTGGAATCCACCTGTTCCAGCAAG 14337
1556 IleGlyThrIleArgSer.HisProSerAlaSerLeuTyrLeuLeuMet 1572
: : : : :
14338 ATGACAAGCTTGAGGGCTCCATCCT.....CTATTCA 14369
1572 rGlnIAsn.PheGluIleGlyAsnAsnPheAlaArgValLysMetGlnVa 1588
: : : : :
14370 ACCGATCTCTCGAA..... 14384
1588 lPrometSerLeuSerSerLeuValGlyThrSerGlnAsnPheAsnGluG 1605
: : : : :
14385 .GCAGTCTCGGTCCAAAGTCTTCTTCAGTATTCAG.....G 14421
1605 lUpheLeuArgArgSerLeuLys.....ThrIleLeuThrTyr 1617
| : : : : |
14422 AAAATTCTCGCGCGGGTTGAACGGTCCGAGCGTGCCCATCAGCACCCCTA 14471
1618 AlaGluGluAsp.....leuGluLe 1624
: : : : :
| : : : : |

14472 CCAC TTCAGGATGGAATCGTCGACCCTACAAGACAGGGCCGTGGATGT 14521
1624 uArgGluThrThrPhePro.....AspGlnValGlnAspLeuValPhea 1639
: : : : :
14522 CCAGAAGACCGGAATATCCTTCGTGATTCCTCTGTGTGATGTGTCCATG 14571
1639 snLeuHisMetIleLeuSerAspThrVal..... 1648
: : : : :
14572 AGCAGTCTCGATCAACCCGATTCATTGTCACATGATGCTCGGAG 14621
1649 LysMetLysGluHisGlnGluAspProGluMetLeuIleAspLeuMetTy 1665
| : : : : |
14622 AAGCTCAGCTACGCCCAGCTCGACAGGGA.....TCTGA 14656
1665 rArgIleAlaLysGlyTyrGlnThrSerProGluArgLeuThrTrpLeuG 1682
| : : : : |
14657 CAGGTCGCTCGC.....TGGCTC. 14675
1682 lAsnMetAlaGlyLysHisSerGluArgSerAsnHisAlaGluAlaAla 1698
: : : : :
14676CGTCACCGTTCTTTCAGCTCCGACACGTTAATCGCG 14711
1699 GlnCysLeuValHisSerAlaAlaLeuValAlaGluTyrLeuSerMetLe 1715
14712 GTGTGGCGCCACGGTCTTGGAGAGCATCATCGCGTTCTCGGAATCCCT 14761
1715 uGlu...AspArgLysTyrLeuProValGlyCysValThrPheGlnAsnI 1731
| : : : : |
14762 CAAGCGAACCCTTGCGTACCTACCTCTG..... 14789
1731 leSerSerAsnValLeuGluGluSerAlaValSerAspAspValValSer 1747
: : : : :
14790GATGTCAGGCTCCTGCTGCCCGCATGTGATCCATGTATCG 14831
1748ProAspGluGluGlyIleCysSerGlyLysTyrPheThrGluSe 1762
| : : : : |
14832 TCGCTACCCGGGACACAAGCTCATCTGCTGGCGCAACGTTACGCCGCC 14881
1762 rGlyLeuValGlyLeuLeuGluGlnAlaAlaAlaSerPheSerMetAlaG 1779
| : : : : |
14882 C.....AAGCTTCAGAAAGCGGCATCGATTTC..... 14909
1779 lYmetTyrGluAlaValAsnGluValTyrLysValLeuIleProIleHis 1795
14910GTGCCCATC... 14918
1796 GluAlaAsnArgAspAlaLysLysLeuSerThrIleHisGlyLysLeuG 1812
| : : : : |
14919CGTGATACCTTCACCTACACTCACT..GACGGCACACTTCA 14956
1812 nGlu..AlaPheSerLysIleValHisGlnSerThrGlyTyrGluArgMe 1828
| : : : : |
14957 AGATGGGCTTAC.....CATCGAGCAGCCCTTCGCCAAGGCC 14994
1828 tPheGlyThrTyrPheArgValGlyPheTyrGlyThrLysPheGlyAspL 1845
: : : : :
14995 TAGCGTACGCCCATGTTCAACGTGTGGTTCTACCGGACGACCGAAGGTTGT 15044
1845 eu.....AspGluGln.GluPheVa 1851
15045 ATGGTCCAGCACCGCAACATGTCGCTTGGTGAAGAAGACGTAACGTGCT 15094
1851 lTyrLysGluProAlaIleThrLysLeuAlaGluIleSerHis 1865
| : : : : |
15095 CGCTAAGCAGCCCGCAGCAGCTCGCATAGCACATATATCCAAAT 15137
seq_name: /cgn2_6/plodata/2/lna/5B_COMB.seq:US-08-457-273B-7
seq_documentation_block:
; Sequence 7, Application US/08457273B
; Patent No. 584995
; GENERAL INFORMATION:
; APPLICANT: Hayden, Michael

2424 2424
372 sasnlsglulysleuGlulysleuYsserGlnAlaaspGlnpheCysg 389
2424 2424
389 lnaArgleuGlyLysTYrArgMetProPheAlaTrpThrAlaIleHisleu 405
2425TGG.....CTG 2430
406 MetAsnIleValSerSerAlaGlySerLeuGluArgAspSerThrGluVa 422
2431 GGCAACATCAGAACCTGCAGAGAAAT..... 2457
422 lGluIleSerThrGlyGluArgLysGlySerTrpSerGluArgAsnS 439
2458ACAT 2461
439 erSerIleValGlyArgSerLeuGluArgThrThrSerGlyAspAsp 455
2462 TTTCCTCTGGTGCATGCATTCCTTACTGCAGAAACGTTGAAGGATGAA 2511
456 Ala.....CysAsnLeu.....ThrSerPheArgProAlaThrLe 467
2512 TCCTTCGTACTGCACAGTTGGCTGTGTACAGCTGTGAGGCACCTGTGTCCT 2561
467 uThrValThrAsnPhePheLysGlnGlyAspArgLeuSerAspGluA 484
2562 GAGTCTTTGCAGC.....AGCAGCTACAGTGACTTGG 2593
484 spleuTYrLysPheLeuAlaaspMetArgArgProSerSerValLeuArg 500
2594 GATTA..... 2598
501 ArgLeuArgProIleThrAlaGlnLeuLysIleaspIleSerProAlaPr 517
2599CAACTGCTTATGTATGCTGCCTCTGAA 2627
517 oGluAsnProHisTYrCysLeuThrProGluLeuGlnValLysLeuT 534
2628 GAACAGCTCCTACTGGCTGGTGAGGACCGAAGCTGCTGCACACTCTGGCAG 2677
534 yfProAspSerArgValArgProThrArgGluIleLeuGluPheProAla 550
2678 AGATTGACTTCAGGGCTC.....GTGAGTTTTTGGAGGCAGCAAAAGCA 2718
551 ArgAspValTYrValProAsnThrThrTYrArgAsnLeuLeuTYrIleTY 567
2719 GAAAGTTTACACCGAGGGCTCATCATATACAGGGTTTCTAAAAATACA 2768
567 rProGlnSerLeuAsnPheAlaAsnArgGlnGlySerAlaArgAsnIleT 584
2769 AGAAGAGTACTCAAT.....A 2785
584 hrValLysValGlnPheMetTYrGlyGluAspPro..... 595
2786 ATGTGGTCATTTATTTGCTTGGAGATGAAGAGACCCAGGGTTCGACATGTT 2835
596SerAsnAlaMetProValIlePheGlyLysSerSe 607
2836 GCTGCAACATCATTAACAAGGCTTGTCCCAAGCTGTTTACAAAG.... 2880
607 rCysSerGluPheSerLysGluAlaTYrThrAlaValAlaTYrHisAsnA 624
2881 .TGTGACCAGGACAGCTGATCCAGTTGTGGCTGTAGCGAGGATCAGA 2929
624 rGserProAspPheHisGluGluIleLysValLysLeuProAlaThrLeu 640
2930 GCAGT.....GTCTACCTGAAGCTCCTCATGCATGAG 2961
641 Thr.AspHisHisIleLeuPheThrPheTYrHisValSerCysGlnG 657
2962 ACCCAGCCACCATCAC..... 2977

657 lnlYsglnAsnThrProLeuGluThrProValGlyTYrThrTrpIlePro 673
2978ACTTTTCTGTGACACCA..... 2995
674 MetLeuGlnAsnGlyArgLeuLysThrGlyGlnPheCysLeuProValSe 690
2995 2995
690 rLeuGluLysProProGlnAlaTYrSerValLeuSerProGluValProL 707
2996TCACCAGAA..... 3004
707 euProGlyMetLysTrpValAspAsnHisLysGlyValPheAsnValGlu 723
3005TCTATAGAG 3013
724 ValValAlaValSerSerIleHisThrGlnAspProTYrLeuAspLysPh 740
3014 GCTATAGCTTACTGCCAAGAATACAGATGTCCACCATGG..... 3052
740 ePheAlaLeuValAsnAlaLeuAspGluHisLeuPheProValArgIleG 757
3053AAACAATCTCTCAAGAGTTGTGCCG..... 3079
757 lYAspMetArgIleMetGluAsnAsnLeuGluAsn..... 768
3080CAGTTTCTCATGAAGCTCATTAAGCAACAACAGGGCACTCAC 3121
769GluLeuLysSerSerIleSerAlaLeuAsnSerSe 780
3122 ATTTGGATGCTGTGAAGCCTTGTCTCTCTCAGCAGCCTTTCAGTTT 3171
780 rGlnLeuGluProValValArgPheLeuHisLeuLeuAspLysLeuI 797
3172 GCACTTGGAG..... 3181
797 lLeuLeuValIleArgProProValIleAlaGlyGlnIleValAsnLeu 813
3182TTTAG.....ATGGCACTG 3196
814 GlnGlnAlaSerPheGluAlaMetAlaSerIleIleAsnArgLeuHisLy 830
3197 TGGAGTGGCCCCACTGAGTGCCTGTGATGAGTCCAGGAGAGCTGCAC.. 3244
830 sasnLeuGluGlyAsnHisAspGlnHisGlyArgAsnSerLeuLeuAla 847
3245 .TGTGGGATGGCCTCCATGATTTCTCAC.....CTTG 3275
847 eTYrIleHisTYrValPheArgLeuProAsnThrTYrProAsnSerSer 863
3276 CTTTCATCAAGCTTGGT.....CCCACTGATCT 3304
864 SerProGlyProGlyGlyLeuGlyGlySerValHisTYrAlaThrMetAl 880
3305 CTCAGCCCATCAGGATGCCCTTGATTTTGGCTGGAACCTTGCCTAGCAGCGA 3354
880 aArgSerAlaValArgProAlaSerLeuAsnLeuAsnArgSerArgSerL 897
3355 GTGCCCCCAAGTCTCTGAGAAGTTTCATGGACCTCTGAAGAAGAAGCCAAC 3404
897 euSerAsn.....SerAsnProAspIleSerGly..... 906
3405 TCAGCAGCCACACAGACAGAGAAATCTGCCCTGCTGTGGGGGATCGGAC 3454
907ThrProHiserProAspAs 913
3455 TCTAGTGCCCTTGGTGGAGCAGCTTTTCTCCACACCTGCTGAAGGTGATCA 3504
913 pGluValArgSerIleIleGlySerLysGlyLeuAspArgSerAsnSerT 930
3505 ATATCTGTGCTCATGTCTTGA.....CGATGTGACTCTCCT 3539

4998 CCGTGGACATGCTTTGGCGAGTATGTTCAATCACTCCAAACAAATGGCA 5047
1487 ValThrGlnSerLysGlnSerIleLeuGlyValLeuLysValLeu 1503
::: ::::: |||
5048 TCTGTAAACACTGTGACCTGTGGATATCTGGAATCCTCGCATTTCTG... 5095
1503 uHisSerMetAlaCysAsnGlnSerAlaValTyrLeuGlnHisCysPhe 1520
::: |||
5096 ...AGGTTCTCATTTCCAGTCA..... 5116
1520 LeThrGlnArgAlaLeuValSerLysPheProGlnLeuLeuPheGlnGlu 1536
|||::: ::::: |||
5117 .ACCGAGGACATTTGTTCTTGTCTGATTCAGAGCTCTCTCTCCA 5164
1537 GluThrGlnGlnCysAlaAspLeuCysLeuArgLeu..... 1548
::: |||
5165 CACTTGCTCTCTGT...CCAGTATTACAGGTTAAGGGGTGAGGCGG 5211
1549LeuArgHisCysSerSerSerIleGlyThrIleArgSerH 1562
|||::: |||
5212 TAATGTAACACTAGGAGATGCAGCGAA.....GGGAAACAAAGAGTT 5255
1562 isProSerAlaSerLeuTyrLeuLeuMetArgGlnAsnPheGluIleGly 1578
|||:::
5256 TGCCAGAA..... 5263
1579 AsnAsnPheAlaArgValLysMetGlnValProMetSerLeuSerIle 1595
::: |||
5264 GATACATTTCTCAAGGTTTCTTTTACAGCTGCTGTGTAATCTTCTAGAGA 5313
1595 uValGlyThrSerGln.....AsnPheAsnGlnGlu.....P 1606
::: |||
5314 CATCGTTACAAACAGCTCAAGTGACATGAGTGAACAGCAGCATACGT 5363
1606 heLeuArgArgSerLeuLysThrIleLeuThr..... 1616
|||::: |||
5364 TCTACTGCCAAGAGCTAGGCACACTGTCATGTGTCTGATCCACATATTTC 5413
1616 1616
5414 AAATCTGGAATGTTCCGGAGATCACAGCAGCTGCCACTAGACTCTTTCAC 5463
1617TyrAlaGlnGluAspLeuGlnLeu 1625
|||::: |||
5464 CAGTGATGGCTGTGAAGGCACTTCTATCTAGAGACCTGAATGCAC 5513
1625 rgGluThrThrPheProAspGlnValGlnAspLeuValPheAsnLeuHis 1641
|||::: |||
5514 GGGTCGATCCATGGTGCCACGCAACCAGCCCTGTA..... 5551
1642 MetIleLeuSerAspThrValLysMetLysGlnHisGlnGluAspProGln 1658
::: ::::: |||
5552 CTGCTGTGGTGCAGATCCTACTTCTCATCAACCACTGAC..... 5593
1658 uMetLeuIleAspLeuMetTyrArgIleAlaLysGlyTyrGlnThrSerP 1675
|||
5594TACCGGTGGTGGCGAGAGGTGCAGCAGACAC 5624
1675 roGluArg.....LeuThrTrpLeuGlnAsnMet.....Ala 1685
|||::: |||
5625 CCAAGAGACACAGTCTGTCTCCSCACGAAAGTCACTTAACCCCAAGATCT 5674
1686 GlyLysHisSerGlnArgSerAsnHisAlaGlnAlaLagLncys.... 1700
|||::: |||
5675 GCGCAAGAGAGGATTTCTGCGCGGCACTCAAGCTGGGAATGTGCAATAG 5724
1701 ...LeuValHisSerAlaAlaLeuVal.....AlaGluTyrLeuSerH 1714
::: |||
5725 AGAATAGTGGCGAGAGGCGCTTATCTCTCTGTGATTTATGTCTGTC 5774
1714 et...LeuGluAspArgLysTyrLeu..... 1721
::: |||
5775 AGAATCTTCATGACTCAGAACACTTAACATGGCTCATTTGTAATCACATTT 5824

1722ProValGlyCysValThrPh 1728
|||
5825 CAAGATCTGATCAGCTTGTCTCATGAGCCCTCCAGTA..... 5860
1728 eGlnAsnIleSerSerAsnValLeuGlnGluSerAlaValSer..... 1742
|||::: |||
5861 .CAAGACTTTATTAGTCCATTCATGTAATTTCTGCAGCTAGTGTCTTT 5909
1743AspAspValValSerProAspGlu 1750
::: |||
5910 TTATCCAGGCAATTCACTCTGCTGTAATAATCTTCAACGCCCAACCACT 5959
1751GluGlyIleCysSerGlyLysTyrPh 1759
|||
5960 CTGAAGAAAACACTTCAGTGTGCTTGAGGCATC.....CATCT 5997
1759 eThrGlnSerGlyLeuValGlyLeuLeuGlu.....GlnAla 1772
::: |||
5998 CAGCCAGTCTGGCGCTGTGCTCACACTTATGTGACAGGCTCCTGGGCA 6047
1772 IaAlaSerPheSerMetAlaGlyMetTyrGlu.....AlaValAsnGlu 1786
::: |||
6048 CCTCTCCCGTGGCGCTGCTGCATGTGTGCACACCCTGGCTGTGCGCGG 6097
1787 ValTyrLysValLeu.....IlePr 1793
|||
6098 GTAGAATGTCTTTGGCTGCAAAATTTCAGACAGCATGGCCCACTTGGC 6147
1793 oIleHisGluAlaAsnArgAspAlaLysLysLeuSerThrIleHisGlyL 1810
|||
6148 AGAGGAGGAACTGAACAGA.....ATCCAAGAAC 6176
1810 ysLeuGlnGluAlaPheSerLysIleValHisGlnSerThrGlyTyrGlu 1826
|||
6177 ACCTCCAGAACAGTGGGCTTGACACAAAGACACAA..... 6211
1827 ArgMetPheGlyThrTyrPheArgValGlyPheTyrGlyThrLysPheGln 1843
|||
6212 AGCTCTATTCACGTCTGCAGAGA.....TTCCG 6240
1843 yAspLeuAspGlnGlnGluPheValTyrLysGluProAlaIleThrLysL 1860
|||
6241 ACTCTTACTGTGCAGGACTCACTTAGCCCTTGCCCCAGTCACT... 6286
1860 euAlaGluIleSerHisArgLeuGlnGlyPheTyrGlyGluArgPheGly 1876
|||
6287TCCACCCACTGGGTGG.....GATGGG 6310
1877 GluAspValValGluValIleLysAspSerAsnProValAspLysCysLys 1893
::: |||
6311 CACACATCTCTGGAAACAGTG..... 6331
1893 sleuAspProAsnLysAla...TyrIleGlnIleThrTyrValGluPro 1909
::: |||
6332 ...AGTCCAGACAAAGACTGTGTAACCTTCAGCTTGTCAAGATCCAGTGT 6377
1909 yr.....PheAspThrTyrGlnMetLysAspArg 1918
::: |||
6378 GGACCCAGATCAGATTCTGCACCTGTGGAAGGTGCAGAGCTGTCAACCT 6427
1919 IleThrTyrPheAspLysAsn.....TyrAsnLe 1928
|||
6428 ATCCCTGCTGAAGATATGAATGACTTCAATGATGAGCTCGGAGTTCAACCT 6477
1928 uArgArgPheMetTyrCysThrProPheThrLeuAspGlyArgAlaHisG 1945
|||
6478 AAGCTTTTGGCTCCCTGTTTAAAGCTTGGCATGAGCGAGATGTCTAATG 6527
1945 lylGluLeuHisGlnGlnPheLysArgLysThrIleLeuThrThrSerHis 1961
|||
6528 GCCAAAAGAGTCCCTCTTTGAAGCA..... 6553

```
1962 AlapheProTyrIleLysThrArgValAsnValThrHisLysGluLuiL 1978
      ::::
6554 .....GCCCCGTGGGT 6564
1978 eileLeuThrProIleGluValAlaIleGluAspMetGlnLysLysThrG 1995
      :|||||::: :::::
6565 GATTCTGAACCGGTGACCAGTGTGTTCAGCAGCTTCCT..... 6604
1995 lngluleuAlapheAlaThrHisGlnAsp.....ProAla 2006
      |||:::|||||
6605 .....GCTGTCCATCAAGTCTTCACGCCCTTCCTGCCTATA 6640
2007 AspProLysMetLeuGlnMetValLeuGlnGlySerValGlyThrTrva 2023
      ::|||
      :|||:::
6641 GAGCCACAGCGCCTACTGGAACAAGTTGAATGATCTGCTTGTAATACCAC 6690
2023 lAsnglnglyProLeuGluVal.....AlaGln.....V 2033
      :::: ||| :::: |||||
6691 ATCATACAGTCTCTGACCATACTTGCCCGCTGCCCTGGCACAGTACCTGG 6740
2033 alpheLeuSerGluIleProSer 2040
      || |||||:::|||||:::
6741 TGGTGCTCTCCAAAGTGCCTGCT 6763

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-328-254-5

seq_documentation_block:
; Sequence 5, Application US/08328254
; Patent No. 5710022
;
; GENERAL INFORMATION:
; APPLICANT: Zhu, Xueliang
; APPLICANT: Lee, Wen-Hwa
; TITLE OF INVENTION: A No. 5710022el Nuclear Mitotic Phosphoprotein
;
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell and Flores
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/328,254
; FILING DATE: 24-OCT-1994
; CLASSIFICATION: 435
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/141,239
; FILING DATE: 22-OCT-1993
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-CJ 1191
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 5:
;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8789 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
;
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 544..7990
;
; US-08-328-254-5
```

```
alignment_scores:
      Quality: 152.50      Length: 1009
      Ratio: 0.330      Gaps: 45
Percent Similarity: 45.788      Percent Identity: 18.831

alignment_block:
US-09-737-246-2 x US-08-328-254-5 ..

Align seg 1/1 to: US-08-328-254-5 from: 1 to: 8789

1022 HisLeuTyrPheAsnAspLysLeuGluAlaProArgLysSerArgPhePr 1038
      ||||| :::::|||||
2821 CACTTGACATTGTTCAGACAAGAAGTTCAAATGCACCTTGGCGAATTGCA 2870
1038 oGluArgPheMetAsp.....AspIleA 1046
      |||:::|||||:::
2871 AGAGAAATTCTTATCTTTACAAAGTGAAACACAAATTTTACATGATCAGC 2920
1046 laAlaLeuValSerThrIleAlaSerAspIleValSerArgPheGln... 1061
      ::::: |||:::
2921 ACTGTGAGATGAGCTCTAAATGTCAAGAGCTGCAGACCTATGTTGACTCA 2970
1062 ..LysAspThrGluMetValGluArgLeuAsnThrSerLeuAlaPhePh 1077
      ||| ::::: ||| |||::: |||
2971 TTAAGGCCGAAATTTGGTC.....TTGTCAACGAATCTGAGAAACTT 3014
1077 eleuAsnAspLeuLeuSerValMet.....AspArgGlyPheV 1090
      | ::||| |||::: ||| ::::: |||:::
3015 TCAAGGTGACTTGGTGAAGGAGATGCAGCTGGGCTTGGAGGAGGGGCTCG 3064
1090 alpheSerLeuIleLysSerCysTyrLysGlnValSerLysLeuTyr 1106
      || ||||| ::||| |||::: |||
3065 TTCCATCCCTGTCACTCCCTGTGTGCTGCTGAC..AGCTTAGTCTTAGC 3111
1107 SerLeuProAsnProSerValLeuValSerLeu.....ArgLeuAspPh 1121
      ||||| ::||| |||::: |||
3112 AGTTGGGAGACTCCTCCTTTTACAGAGCTCTTTAGACAGACAGAGAGA 3161
1121 eleuArgIleIleCysSerHisGlnHisTyrValThrLeuAsn..... 1135
      ::::: ::::: ||| |||::: |||
3162 TATGTCTCTTTTGTAGTAATTTAGAAAGGGGCTGTTTCAGCAAAACAGTGCA 3211
1136 .....LeuProCysSerLeuLeu..... 1141
      ::::: ||||| |||
3212 GTGTAGATGAAGTATTTTGCAGCAGCTCTGCAGGAGAGAATCTGACCAGG 3261
1142 ..ThrProAlaSerProSerProSerValSerSerAlaThrSerG1 1157
      |||::: |||::: |||::: |||
3262 AAAGAAACCCCTTGGCCCCCAGCGAAGGGTGTGTAAGAGCTTGAGTCCCT 3311
1157 nSerSerGlyPheSerThrAsnValGlnAspGlnLysIleAlaAsnMetP 1174
      ::::: :::::
3312 CTGTGAGGTGTACCGGCAGTCCCTCGAG..... 3339
1174 heGluLeuSerValProPheArgGlnGlnHisTyrLeuAlaGlyLeuVal 1190
      ::|||::: ::::: |||
3340 ..AAGCTAGAGAGAAATGGAAGTCAAGGATTTATGAATAATAAGGAA 3387
1191 LeuThrGluLeuAlaValIleLeuAspProAspAlaGluGlyLeuPheG1 1207
      ::||| |||::: |||::: |||
3388 ATTCAGAGAGCTCGAGCAGTATTTAAGTTCTGAAGAAGCAAGAGCTTGACTG 3437
1207 yLeuHisLys.....LysValI 1213
      |||::: |||
3438 CCTTAGAGAGCATATTTTGTACAGAAATGCAACAGTGGCAACAGAGCTGA 3487
1213 leAsnMetValHisAsnLeuLeuSerSerHisAspSerAspProArgTyr 1229
      ::::: |||::: |||::: |||
3488 CAAGCGTGACTCTGAGATGAGTCCAGTGTGGCGGAGAAAGAAACAG 3537
1230 SerAspProGlnIleLysAlaArgValAlaMetLeuTyrLeuProLeuI1 1246
      ::::: |||::: |||::: |||
3538 ACGGAA...CACTGTCACTTGAGCTGGAAGTAGCACGAGCTCCAGCTTACA 3584
```

1246 eglylleilemetglutThrValProGlnLeuTyraspheThrGluThrH 1263
|||::: :|||
3585 AGGCTTGAGACTTAAGTCTCGGCTCTTGGCTTGGCATCGACACAGAGATG 3634
1263 isanGlnArgGlyArgProIleCysIleAlaThrAspAspTyrGluSer 1279
||| :|||
3635 CTATTCAAGGCCGAATGAGAGCTGTGACATATCAAAAGACATACTTCA 3684
1280 GluSerGlySer.....MetIleSerGlnTh 1288
|||::: :|||
3685 GAAACTACAGAAAGAACACCAAGCATGATGTTCAATGAGATTGTGATTA 3734
1288 rValAlaMetAlaIleAlaGlyThrSerValProGlnLeuThrArgProg 1305
|||::: :|||
3735 AGATGCTCAGCAGAGACCTCAATCTAGACATTGAGAAATAACTGAGACTG 3784
1305 lySerPheLeuLeuThr...SerThrSerGlyArgGlnHisThrThrPhe 1320
|||::: :|||
3785 GTGCATTGAAACCACAGAGAGAGTGCTCTGGGGAACAGTCCCCAGATACC 3834
1321 SerAlaGlu.....SerSerArgSerLeuLeuIleCysIle 1332
::: ||| :|||
3835 AATTATGAGCCTCCAGGGGAGAGATAAACCAGGCTCTTCAGAAATGCAT 3884
1332 uLeuTrpValLeuLysAsnAlaAspGluThrValLeuGlnLysTrpPhe. 1348
: :||| :|||
3885 TTCTGAATTGTCATTCTGCTGCTCAATGCTTGCTACCTATGATTTCC 3934
1349ThrAspLeuSerValLeuGlnLeu..... 1356
|||::: :|||
3935 TGGGGAATCAGGAAGATATCCATAATCTCAACTGCGGGTAAAGAGACA 3984
1357 ..AsnArgLeuLeuAspLeuLeuTyrLeuCysValSerCysPheGluTyr 1372
|||::: :|||
3985 TCAATGAGAAATTGAGATTACTTCAATG.....ATAGAGGA 4022
1372 rLysGlyLysLysValPheGluArgMetAsnSerLeuThrPheLysLys 1389
::: :||| :|||
4023 CCGTGACAGAAAAGTTGAAAGTTTGCTAAATGAAATGAAAGATTAGACT 4072
1389 erLys.....AspMetArgAlaLysLeuGlu..... 1397
||| :|||
4073 CAAAACCTCCATTACAGAGAGGTACAACTAATGACCAAATTTGAAGCATGC 4122
1398GluAlaIleLeuGlySerIleGlyAla..... 1406
||| :|||
4123 ATAGAATTGGAAAAAATAGTTGGGGAACCTTAAGAAAGAAACTCAGATTT 4172
1407ArgGlnGluMetValArgA 1413
4173 AAGTGAAAAATTGCAATATTTTCTTGATGATCACCAGAGTTACTCCAGA 4222
1413 rg.....SerArgGly.....GlnLeuGluArgSerProSer 1423
|| :||| :|||
4223 GAGTAGAAACTTCTGAAAGCCTCAATTCCTGATTTAGAAATGCATGCAGAT 4272
1424 GlySerAlaPheGlySerGlnGluAsnLeu..... 1433
|| :||| :|||
4273 AAATCA.....TCACGTGAAGATATGAGATAATGTGCCCAAGT 4313
1434ArgTrpArgLysAspMetThrHisTrpArgGlnAsnThrGluL 1448
|||::: :|||
4314 GAATGACAGCTGGAAGAGAGAGATTTCTTGATGTGAAAAATGAGCTGAGTA 4363
1448 ysLeuAspLysSerArgAlaGluIleGluHisGluAlaLeuIleAspGly 1464
::: :||| :|||
4364 GGATCAGATCGGAGAAAGCTAGCATTTGAGCATGAAGCCCTCTTAC..... 4407
1465 AsnLeuAlaThrGluAlaAsnLeuIleIleLeuAspThrLeuGluIleVa 1481
|||::: :|||
4408CTGAGGCTGACTTAGAGGTAGTTCAAAACAGAGAAAGCTATG 4448

1481 lVal.....G 1483
::: :|||
4449 TTAGAAAAAGACAATGAAATTAAGCAGAAGCTTATTGTCTGCCCTTGAAG 4498
1483 lNThrValSerValThrGluSerLysGluSerIleLeuGlyGlyValLeu 1499
: :||| :|||
4499 AAGAACTCTCAGTGGTCCACAAGTAGAGAAACACGCTTCTGGAAGATTTA 4548
1500 LysValLeu..... 1502
::: :|||
4549 GATACTATGTCAAAAAAACACGCGCAGCTGATCAGTTGTGAAAAAAT 4598
1503LeuHisSerMetAlaCysAsnGlns 1511
|||::: :|||
4599 GAAGAGAAACACACAAGAGCTTGAGTCTCATCAAAGTGAAGT..... 4641
1511 erAlaValTyrLeuGlnHisCysPheAlaThrGlnArgAlaLeuValSer 1527
|||::: :|||
4642CTCCATTGCATTCAGGTGGCAGAGAGCAGAGTG... 4674
1528 LysPheProGluLeuLeuPheGluGluGluThrGluGlnCysAlaAspLe 1544
::: :||| :|||
4675AAGAAAAAGACGAA..... 4689
1544 uCysLeuArgLeuLeuArgHisCysSerSerSerIleGlyThrIle... 1559
|||::: :|||
4690CTCCTTCAGACTTTGTCTCTGATGTGAGTGAAGCTGTTAA 4729
1560ArgSerHisProSerAlaSerLeuTyrLeuLeuMetArgGlnAsn 1574
::: :||| :|||
4730 AAGACAAAACTCATCTCCAGGAAAGCTGCAGAGTTTGAAAAAGACTCA 4779
1575PheGluIleGlyAsnAsnPheAlaArgVa 1584
|||::: :|||
4780 CAGGCACTGCTTTTGACAAAAATGTGAGCTGGAACCAAAATTCACACACT 4829
1584 l..... 1584
: :|||
4830 GAATTAAGAGAAAGAATTGCTTGTCAAGGAATCTGAAGCCTGCAGGCCA 4879
1585LysMetGlnValPrometSerLeuSer 1593
|||::: :|||
4880 GACTGAGTGAATCAGATTATGAAAAAGCTGAATGTCTCCAAAGGCTTGAG 4929
1594 Ser.....LeuValGlyThrSe 1599
::: :||| :|||
4930 GCCGCACTGGTGGAGAAAGGTGAGTTCGCATTGAGGCTGAGCTCAACACA 4979
1599 rgLAsnPheAsnGlnGluPheLeuArgArgSerLeuLysThrIleLeuT 1616
::: :||| :|||
4980 GGAGGAAGTGCATCAG.....CTGAGAAAGAGCATCGAGAAACTGAGAG 5023
1616 hrTyrAlaGluGluAsp.....Leu 1622
: :||| :|||
5024 TTGCGCATTTGAGGCGGATGAAAGAAAGACAGCTGCACATCGCAGAGAACTG 5073
1623 GluLeuArgGlu.....ThrThrPheProAspGlnValGlnAspLe 1636
::: :||| :|||
5074 AAAGAAACGCGAGCGGAGAAATGATTCACCTTAAGGATAAAGTTGAGAACCT 5123
1636 uValPheAsnLeuHisMet.....IleLeuSerA 1646
: :||| :|||
5124 TGAAGAGGAATTGCAGATGTCAAGAAACACAGAGCTTAGTATTCTTG 5173
1646 spThrValLysMetLysGlnHisGlnGluAspProGluMetLeuIleAsp 1662
|||::: :||| :|||
5174 ATGCCGAGAATTCCAAAGCAGAAAGTAGAGACTTAAAAAAACAAATAGAA 5223
1663 LeuMetTyrArgIleAlaLysGlyTyrGln..... 1672
||| :||| :|||
5224 GAGATGGCCAGAAAGCCTGAAGTTTGTGAATTAGACCTTGTACAGTTAAG 5273
1673 .ThrSerProGluArgLeuThr...TriLeuGlnAsnMetAlaGlyLysH 1688

```

::: ||::| ||| ::||::| ||::|
5274 GTCTGAAAGAAATCTGACAAACAAATACAGAAACAAAGTCAGT 5323
1688 lSserGluArgSerAsnHisAlaGluAlaAlaGlnCysLeuValHisSer 1704
5324 TGTCAAGACTAGACAAG..... 5340
1705 AlaAlaLeuValAlaGluTyrLeuSerMetLeuGluAspArgLysTyrLe 1721
5341 .....TFACTCTCTTCATTTAAGTCTGTAGAGAAAGAGCAAGC 5384
1721 uProValGlyCysValThrPheGlnAsnIleSerSerAsnValLeuGluG 1738
5385 AGAGATA.....CAGATCAAGAAG 5404
1738 lUserAlaValSerAspAspValValSerProAspGluGluGlyLeCys 1754
5405 AATCTAAACTGCAGTGGAGATGCTT..... 5430
1755 SerGlyLysTyrPheThrGluSerGlyLeuValGlyLeuGluGlnAl 1771
5431 .....CAGAACTCACTTAAAGAGAGCTAAATGAGGCAGT 5462
1771 aAlaAlaSerPheSerMetAlaGlyMetTyrGluAlaValAsnGluValT 1788
5463 AGCAGCC.....TTGTGTGTGACCAGAAATTATGAGGCCACAG 5503
1788 yLysValLeu...IleProIleHisGluAlaAsnArgAspAlaLysLys 1803
5504 AACAGAGTCTAGACCCACCAATAGAGAGAAGCATCAGCTGAGAAATAGC 5553
1804 LeuSerThrIleHisGlyLysLeuGln 1812
5554 ATGAAAGCTGAGAGCCCGCCTAGAA 5580
```

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:us-09-307-143-5

seq_documentation_block:
; Sequence 5, Application US/09307143
; Patent No. 6335157
; GENERAL INFORMATION:
; APPLICANT: Gonzalez C.
; APPLICANT: Lange, B.
; TITLE OF INVENTION: METHODS BASED ON LOCALIZATION OF HSP90 TO THE
; FILE REFERENCE: 9882-003
; CURRENT APPLICATION NUMBER: US/09/307,143
; CURRENT FILING DATE: 1999-05-07
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 2543
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (85)..(2256)
US-09-307-143-5

alignment_scores:
Quality: 150.50 Length: 573
Ratio: 0.593 Gaps: 33
Percent Similarity: 44.328 Percent Identity: 22.339

alignment_block:

US-09-737-246-2 x US-09-307-143-5 ..

Align seg 1/1 to: US-09-307-143-5 from: 1 to: 2543

```

108 GluMetAspPro.HisValArgAspCysIleArgSerTyrThrGluAspT 124
::: ||||| ||::| ||::| ||::| ||::|
621 CAAAGTGATCCTCCATCTTAAAGAAAGATCAGACAGAGTACCTAGAAAGA 670
```

```

124 rPalAlaLe.....ValIleArgLysTyrHisLysLeuGlyThrGlyPhe 138
::: ||::| ||::| ||::| ||::| ||::|
671 GCGGGGTCAAGAGTAGTAGAAGAG...CATTCTCAGTTTCATAGGCTTAT 717
139 AsnProAsnThrLeuAspLysGlnLysGluArgGlnLysGlyLeuProLy 155
718 ..CCCATCACCCCTTATTGAGAGAGAACGAGAGAGGAAATT.... 759
155 sGlnValPheGluSerAspGluAlaProAsp.....GlyAsnSerTyrG 170
760 .....AGTGATGATGAGGCAGAGAGAGAAAGGTGAGAAAGAG 799
170 lAspAspGlnAspAspLeuLysArgArgSerMetSerIleAspAspThr 186
800 AGGAAGATAAAGATGATGAAGAAAG..... 825
187 ProArgGlySerThrPalAcysSerIlePheAspLeuLysAsnSerLeuPr 203
825 ..... 825
203 oAspAlaLeuLeuProAsnLeuLeuAspArgThrProAsnGluGluIleA 220
826 .....CCCAAGATCGAAGATGTGGCTTCAGATGAGAGGATG 862
220 sPArgGlnAsnAspAspGlnArgLysSerAsnArgHisLysGluLeuPhe 236
863 ACAGCGGTAAAGGATTAAGAAAGAAACTAAGAGATCAAAAGAGAAATAC 912
237 AlaLeuHisProSerProAspGluGluProIleGluArgLeuSerVa 253
913 ATTGATCAGGAAGAACTAAACAAGACCAGCCTATTGACCAGAAACC 962
253 lProAspIleProLysGluHisPheGlyGln..... 263
963 TGATGACATCACCCAGAGAGAGTATGAGAAATCTCAAGAGCCTCACTA 1012
264 .....ArgLeuLeuValLysCysLeuSerLeuLysPheGlu 275
1013 ATGACTGGGAAGACCACTTGGCAGTCAAGCACCTTTCTGTAAGAGTCAAG 1062
276 lIleGluIleGlu.....ProIlePheAlaSerLeuAlaLe 287
1063 TTGGAATTCAGGGCACTTGTATTATTCCCTCGCGGCTCCCTTTGACCT 1112
287 uTyrAspValLysGluLysLysLys.....IleSerGluA 299
1113 TTTTGAGAACAGAGAAAGAAACAACATCAACTCTATGTCCGCCGTG 1162
299 snPheTyrPheAspLeuAsnSerGluGlnMetLysGlyLeuLeuArgPro 315
1163 TGTTCATCATGACAGCTGTGATGAGTTGATACAGAGTATCTC..... 1206
316 HisValProProAlaAlaIleThrThrLeuAlaArgSerAlaIlePheSe 332
1207 .....AATTTCCTCGTGGTGTGACTTC 1232
332 rIleThrTyrPro.....SerGlnAspValPheLeuValIleLysL 346
1233 TGAGGATCTGCCCTGACACATCTCCGAGAAATGCTCCAGACAGCAAA 1282
346 euGluLysValLeuGlnGlnGlyAspIleGlyLysCys..... 358
1283 TCTTGAAGATCATTCGCAAAACATTTGTAAGAGTGCCCTTGAGCTCTTC 1332
359 .....AlaGluProTyrMetIlePheLysGluAla.. 368
1333 TCTGAGCTGCGAGAGACAGAGAAATTTACAAGAAATTTCTATGAGGCATT 1382
369 .....AspAlaThrLysAsnLysG 375
1383 CTCTAAAAATCTCAAGCTTGAATCCACGAGAGACTCCACT..AACCGCC 1429
```


[illegible]

```

seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-629-001A-2
seq_documentation_block:
; Sequence 2, Application US/08629001A
; Patent No. 5858661
; GENERAL INFORMATION:
; APPLICANT: Shiloh, Yosef
; TITLE OF INVENTION: ATAXIA-TELANGIECTASIA GENE AND ITS
; TITLE OF INVENTION: GENOMIC ORGANIZATION
;

```

```

: NUMBER OF SEQUENCES: 139
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Kohn & Associates
: STREET: 30500 No. 5858661thwestern R
: CITY: Farmington Hills
: STATE: Michigan
: COUNTRY: US
: ZIP: 48334
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Ve
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/629, 001A
: FILING DATE:
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Kohn, Kenneth I.
: REGISTRATION NUMBER: 30,955
: REFERENCE/DOCKET NUMBER: 2290.00032
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (810) 539-5050
: TELEFAX: (810) 539-5055
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 9171 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: ORIGINAL SOURCE:
: ORGANISM: Homo sapiens
: POSITION IN GENOME:
: CHROMOSOME/SEGMENT: 11q22-23
US-08-629-001A-2

```

[illegible]

```
alignment_block:
US-09-737-246-2 x US-08-629-001A-2  ..
```

Align seg 1/1 to: US-08-629-001A-2 from: 1 to: 9171

267 vallyscysleuSerLeuLysphecgluilegluilegluProtlephea 283
||| |||||::||| ::::||| ||| :::
1615 GTATGCTGTTTGACTTTGGCACCTGACCACCAGTATAGTTCACAGGACGCT 1664
283 aSerleuAlaleuTyrrAspVallysgLylsLyslyleSerGluAsn 300
:::||||:::| ||:::| ::::| |||
1665 AAAAATGGGAATA.....GAGCAAAATATGTGTGAAGTAATA 1702
300 heTyrrPheaspLeuAsnSerGluGluMetlysglyleu..... 313
|||::|||:::|:::| ||||| |||||
1703 GAAGCTTTCTTTAAAGGAATCAATTAATGAATGGCTCTTATTCTATCAG 1752
314ArgProHisValProProAlaAlaIleTh 323
:::|||||||
1753 TTAGAGGCTGACTTAGAAAATAGCACAGAGTGCCCTCA..... 1791
323 rThrleuAlaArgSeraAlaIlePheSerIleThTyrrProSerGluAsp 340
|||::| ::::| |||
1792ATTCCTTCACAGTAATTTTCCCTCAT..... 1815
340 alPheleuValIleLysleuGluLysValleuGlnGlnGlyaspIleGly 356
:::| |||||:::| ||| ::::|
1816CTTGTACTGGAGAAATTCCTGTGAGTCTCAGCTATGAAA 1854
357 GluCysAlaGluProTyrrMetIlePheLysGluAlaAspAlaThrLysAs 373


```
4818 TCCATTGACAAGACTTGAAGACTAAAGATCTTCGAAGACACACTGGAAC 4867
      :          :          |||||:          :          :
1398 .....GluAlaIleLeuGlySerIleGlyAlaArgGlnGlu..... 1409
      :          :          :          :          :          :
4868 TACATAAGATCAGATGGTGGACATTATGAGAGCTTCTCAGGATATCCG 4917
      :          :          :          :          :          :
1410 .....MetValArgArgSerArgGlyGlnLeuGluArgSerPr 1422
      :          :          :          :          :          :
4918 CAAGATGGGATTATGGTGAAACTAGTTGTCAATTGTTCAGATTATCCA 4967
      :          :          :          :          :          :
1422 oSer.....GlyS 1425
      :          :          :          :          :          :
4968 GATGGCAATAAACCACTGGTGAAGAAAGTCTAGAGGCTGTGGAA 5017
      :          :          :          :          :          :
1425 eraIaPheGlySerGlnGluAsnLeuArgTrpArgLysAspMetHis 1441
      :          :          :          :          :          :
5018 GCTGCTTGGGA..... 5028
      :          :          :          :          :          :
1442 TrpArgGlnAsnThrGluLysLeuAspLysSerArgAlaGluIleGluHi 1458
      :          :          :          :          :          :
5029 .....GAGTGGGCTCCTATAGATTTCCTACCATAGCTATACACA 5069
      :          :          :          :          :          :
1458 sGluAlaLeuIleAspGlyAsnLeuAlaThrGluAlaAsnLeu..... 1472
      :          :          :          :          :          :
5070 TAGTAA.....GATGCATCTTATACCAAGGCCCTTAAGTTATTGGAAG 5113
      :          :          :          :          :          :
1473 .....IleIleLeuAspThrLeuGlu..... 1479
      :          :          :          :          :          :
5114 ATAAAGAACTTCAGTGGACCTTCATATGCTGACCTACCTGAATTAACACA 5163
      :          :          :          :          :          :
1480 ileValIleGlnThrValSerValThrGluSerLysGluSerIleLeuG1 1496
      :          :          :          :          :          :
5164 CTGGTAGAAGATTGTCTCAAGTTCGATCAGCAGCTGTACCTGTTTGA 5213
      :          :          :          :          :          :
1496 yGlyValLeuLysValLeu.....LeuHisSerM 1506
      :          :          :          :          :          :
5214 AAACATTTTAGCCACAAGACTGGACATAGTTTCTGGAGATTATTAAGA 5263
      :          :          :          :          :          :
1506 etAlaCysAsnGlnSerAlaValTyrLeuGlnHisCysPheAlaThrGln 1522
      :          :          :          :          :          :
5264 TGACACACAGATCCAAATGCTGGCCTATCTACAGCCT..TTTGAACATCA 5310
      :          :          :          :          :          :
1523 ArgAlaLeuValSerLysPheProGluLeuLeuPheGluGluThrG1 1539
      :          :          :          :          :          :
5311 AGAAAAAGTTTTAGAAGTACCCAGATTTGACAAAGAAACCCTTTGA 5360
      :          :          :          :          :          :
1539 uGlnCysAlaAspLeuCysLeuArgLeuLeuArgHisCysSerSerSerI 1556
      :          :          :          :          :          :
5361 AGGCCTGGATGATATA..... 5376
      :          :          :          :          :          :
1556 leGlyThrIleArgSerHisProSerAlaSerLeuTyrLeuLeuMetArg 1572
      :          :          :          :          :          :
5377 .....AATCTGTGATTCCTTAAGT 5397
      :          :          :          :          :          :
1573 GlnAsnPheGluIleGly.....AsnAsnPh 1581
      :          :          :          :          :          :
5398 GAAATCATGACATTTGGATAAGACACTGACTTGTGCTTTTGGACAG 5447
      :          :          :          :          :          :
1581 eAlaArgValLysMetGlnValPrometSerLeuSerSerLeuValGlyT 1598
      :          :          :          :          :          :
5448 TGGAGGCACAAAAATGTGAATTTCTCAATTATTAAAGCCAATGTGTGAAG 5497
      :          :          :          :          :          :
1598 hrSerGlnAsnPheAsnGluGluPheLeuArgArgSerLeuLysThrIle 1614
      :          :          :          :          :          :
5498 TGAATACTGACTTT.....TGTCAGACTGTA 5523
      :          :          :          :          :          :
1615 LeuThrTyrAlaGluGluAspLeuGluLeuArgGluThrThr..... 1628
      :          :          :          :          :          :
5524 CTTCATACTGTGATTCATGATATTTTACTCCAAGATACAAATGAATCATG 5573
      :          :          :          :          :          :
1629 .....PheProAspGlnValGlnAspLeuValPheAsnLeuHisM 1642
      :          :          :          :          :          :
```

```
5574 GAGAAATCTGCTTCTTACACATGTTTCAGGAGTTTTCACACAGCTGTCTTC 5623
      :          :          :          :          :          :
1642 etIleLeuSerAspThrValLysMet..... 1650
      :          :          :          :          :          :
5624 GACACTTCTCGCAAAACGAGCCGATCCACAAACCCCTGCAAACTTGGATTCA 5673
      :          :          :          :          :          :
1651 .....LysGluHisGlnGluAs 1656
      :          :          :          :          :          :
5674 GAGTCAGAGACACTTTTCCGATGCTGTGGATTAATAATCACAAAGAAGAC 5723
      :          :          :          :          :          :
1656 prroGluMetLeuIleAspLeuMetTyrArgIleAlaLysGlyTyrGlnT 1673
      :          :          :          :          :          :
5724 AATGCTTGCTGTGTGGACTACATGAGAAGA.....CAA 5758
      :          :          :          :          :          :
1673 hrSerProGluArgLeuThr.....TrpLeuGlnAsn 1683
      :          :          :          :          :          :
5759 AGAGACCTTCTTCAGGAACAATTTTAATGATGCTTTCGTGGCTGGAT... 5805
      :          :          :          :          :          :
1684 MetAlaGlyLysHisSerGluArgSerAsnHisAlaGluAlaGln.. 1699
      :          :          :          :          :          :
5806 .....TTAAATTATGTAGAAGTTGCCAAGGT 5831
      :          :          :          :          :          :
1700 .....CysLeuValHisSerAlaAlaLeuValAlaGluTyrLeus 1713
      :          :          :          :          :          :
5832 AGCTCAGTCTGTGCTGCTCAGCTTTACAGCTTTACTC....TATGCAG 5875
      :          :          :          :          :          :
1713 erMetLeuGluAspArgLysTyrLeuPro..... 1722
      :          :          :          :          :          :
5876 AATCTATGCAGTAAAGAAAGTATGGATGATCAAGAAAGAAAGTCTT 5925
      :          :          :          :          :          :
1723 .....ValGlyCysValThrPheGlnAsnIleSerSerAs 1734
      :          :          :          :          :          :
5926 GCATTTGAAGAGGAAGCCAGAGTACACAACTATTCTTAGCTTGAGTGA 5975
      :          :          :          :          :          :
1734 nValLeuGluGluSerAlaValSer..... 1742
      :          :          :          :          :          :
5976 AAGTAAAGAAAGAACTGGAATTAAGTTTACAGAGTCTTCTTAGAAATCT 6025
      :          :          :          :          :          :
6026 ACAGAAAGTATAGGGGAGCCAGATAGTTTGTATGGCTGTGGAGGAGAG 6075
      :          :          :          :          :          :
1743 ..AspAspValValSerProAspGluGluGlyIleCysSerGly..... 1756
      :          :          :          :          :          :
6026 ACAGAAAGTATAGGGGAGCCAGATAGTTTGTATGGCTGTGGAGGAGAG 6075
      :          :          :          :          :          :
1757 .....LysTyrPheThrGluSerGlyLeuVal 1765
      :          :          :          :          :          :
6076 ATGTTACAACCCATTACTAGACTACGACATATGAACACGAAAGCAATGTG 6125
      :          :          :          :          :          :
1765 lGlyLeu.....LeuGluGlnAla..AlaAlaSerP 1775
      :          :          :          :          :          :
6126 GGGCAAAAGCCCTAGTACATATGACCTCGAAACAGCAATCCCTCATGCA 6175
      :          :          :          :          :          :
1775 heserMetAlaGlyMetTyrGluAlaValAsn..... 1785
      :          :          :          :          :          :
6176 CACGCCAGGCAAGAAATCATTCAGGCCCTTGCAAGATTGGGACTGTGCCAT 6225
      :          :          :          :          :          :
1786 .....GluValTyrLysValLeuIleProIleHisGluAlaAsnArgAs 1800
      :          :          :          :          :          :
6226 ATTCTTTCCGTCAT.....TTAAAGAGATTGATTAATAATAAAGA 6269
      :          :          :          :          :          :
1800 p...AlaLysLysLeuSerThrIleHisGlyLysLeuGlnGluAlaPhe. 1815
      :          :          :          :          :          :
6270 CTGGTGTCTCGACTAGAGAAGAACTTCAT.....TACCAAGCAGCATGGA 6313
      :          :          :          :          :          :
1816 .....SerLysIleValHisGln 1821
      :          :          :          :          :          :
6314 GGAATATGCAAGTGGGACCAATTGCACCTCCGTCAGCAAAAGAGTAGAAGA 6363
      :          :          :          :          :          :
1822 SerThrGlyTyrGluArgMetPheGlyThrTyr..... 1832
      :          :          :          :          :          :
6364 ACCAGTTACCATGAATCATGTGTACAATGCTGTACAATCTGTAAAGACAG 6413
      :          :          :          :          :          :
1833 ...PheArgValGlyPheTyrGlyThrLysPheGlyAspLeuAspGluG 1848
      :          :          :          :          :          :
6414 AGAATTCTCTACATTTTATGAAGAGTCTCAAAATATGCCAGAGTAAAGAGAG 6463
      :          :          :          :          :          :
```


1848	IngluPheValTyrLysGlu.....Pro	1855
	:::: :::	
6464	TGGAAGAGATGTGTAAAGCGCAGCCTTGAGTCTGTATTCGCTATCC	6513
1856	AlaIleThrLysLeuAlaIuIleSerHisArgLeuGluGlyPheTyrG1	1872
	:::::::::: ::: :: ::: ::	
6514	ACACTTACGAGGTTCAGGCCATTGGA...GAGCTGGAAGCATT...GG	6557
1872	YgluArgPheGlygluAspValVal.....	1880
	::: :: :::	
6558	GGAGCTTTTCTCAAGATCATGACACACATAGACAACTCTCTGAATATATA	6607
1881GluValIleLysAspSer.....	1886
	:::::	
6608	TTAAGTGGCAGAAACACTCCACAGCTTCTCAAGACAGATTTTAACTTTT	6657
1886	1886
6658	CAGAGCCTATCATGGCTCTACGCACAGTCATTTTGGAGATCCTGATGGA	6707
1887AsnProValAspLysCysLysLeuAsp...ProAsnL	1898
	:: ::::	
6708	AAAGGAATGGACAACCTCAACAAGAGAATGTATTAAAGACATTTCTCACCA	6757
1898	ysAlaTyrIleGlnIleThrTyrValGluProTyrPheAspThrTyrGlu	1914
	::::: ::::: :::: :::: ::::	
6758	AACACCTGTGAGAACTCTCTATACTGGCCAGAACTTTCAGAACAACACTCAG	6807
1915	MetLysAspArgIleThrTyrPheAspLysAsnTyrAsn	1927
	::: :: :::: :::	
6808	CTCCCTGAAGGGCAATATTTCAAATTAAACAGATACAT	6846

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-642-274D-2

```

seq_documentation_block:
; Sequence 2, Application US/08642274D
; Patent No. 6200749
; GENERAL INFORMATION:
; APPLICANT: Shiloh, Yosef
; TITLE OF INVENTION: MUTATED FORMS OF THE ATAXIA-TELANGIECTASIA GENE AND METHOD TO
; TITLE OF INVENTION: SCREEN FOR A PARTIAL A-T PHENOTYPE
; FILE REFERENCE: 229000033
; CURRENT APPLICATION NUMBER: US/08/642, 274D
; CURRENT FILING DATE: 1996-05-03
; NUMBER OF SEQ ID NOS: 220
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 9171
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; OTHER INFORMATION:
US-08-642-274D-2

```

```
alignment_scores:
```

Quality:	150.50	Length:	2063
Ratio:	0.175	Gaps:	99
Percent Similarity:	41.638	Percent Identity:	17.402

alignment_block:

US-09-737-246-2 x US-08-642-274D-2

Align seg 1/1 to: US-08-642-274D-2 from: 1 to: 9171

267 ValLYcScysLeuSerLeuLYsPheGluIleGluIleGluProIlePheAl 283
 ||| |||||:::|||| ::||| ||| ::
 1615 GTATGCTGTTTGACTTTGGCACTGACCACCACTAGTATGTTCCAGGAACGCT 1664
 283 aSerLeuAlaLeuTyrAspValLYsGluLYsLYsIleSerGluAsp 300
 ||| |||||:::|||| ::||| ||| ::
 1665 AAAAATCGGAATA.....GAGCAAAATATGCTGTAAGTAATA 1702

[illegible]

421	uValGlulIleSerThrGlyGluArgLysGlySerTrpSerGluArgArgA	438
::::	::: ::: :::::	
2031	TATTTGGCTTCTCTGTCTCCACCAGAACTCTC.....	2058
438	sNserSerIleValGlyArgArgSerLeuGluArgThrThrSerGlyAsp	454
2059AAGGAATCACTGGATCGCTCTTCTGGGA...	2088
455	AspAlaCysAsnLeuThrSerPheArgProAlaThrLeuThrValThrAs	471
2088	2088
471	nPhePheLysGlnGluGlyAspArgLeuSerAspGluAspLeuTyrLysP	488
2089TTATCAGAACACGCTCTCAATAATT	2113
488	heLeuAlaAspMetArgArgProSerSerValLeuArgArgLeuArgPro	504
2114	ACTCATCTGAGATTACAATTCAGAAACTCTTGTCCTCCGGTGTTCACGCTCTT	2163
505	IleThrAlaGlnLeuLysIleAspIleSerProAlaProGluAsnProHi	521
2164	TTGGTGGGTGTCCTGGCTGC.....	2184
521	sTyrCysLeuThrProGlnLeuLeuGlnValLysLeuTyrProAspSera	538
2185	.TACTGTTACATGGGTGTATAAGCTGAAGAGAGAAACATATAAGTCAGAAAT	2233
538	rgValArgProThrArgGluIleLeuGluPheProAlaArgAspValTyr	554
2234	TATTCAGAAAGCCAAAGTCTCTAATGCAATGTGCAGAGAAAGTATC...	2280
555	ValProAsnThrThrTyrArgAsnLeuLeuTyrIleTyrProGlnSerLe	571
2281ACTCTGTTTAAAAAT.....AAGAC	2300
571	uAsnPheAlaAsnArgGlnGlySerAlaArgAsnIleThrValLysValG	588
2301	AAATGAGGAATTCAGAAATTGGTCTTCCTTGAGAAATATG.....ATGC	2341

3767 TTCACATCTGTGATTAGAGTCATTTTGATGAGTGAAGTCCATTGCT 3816
1099
3817 AATCAGATTCAAGAGGACTGAAAAGTCTTCTACAGACTGCTTTCCA
1099
3867 GATTCTTGTAATATCTTCCTTATTTTGCCCTATGAGGGTACCAGAGACA 3916
1100GlnValSerSerLysLeuTyrSerLeuPro 1109
3917 GTGGGATGGCACAGCAAGAGAGACTGTACCAAGCTCTATGATATGCTT 3966
1110 AsnProSerValLeuValSerLeuArgLeuAspPheLeuArgIleIecy 1126
3967 AAAAGTGAAAACTTATGGGAAAAACAGATTGAT..... 3999
1126 sSerHisGluHisTyrValThrLeuAsnLeuPro.....CysSerL 1140
4000CACTTATTCATTAGTAATTACCAGAGATTGTGTGGAGT 4039
1140 euleuThrProProAlaSerProSerProSerValSerSerAlaThrSer 1156
4040 TATTGATGACGTTACATGAGCCAGCAAAATTCTAGTCCAGTCAGAGCACT 4089
1157 GlnSerSerGlyPheSerThrAsnValGlnAspGlnLysIleAlaAsnMe 1173
4090 GACCTCTGTGACTTTTCAGGGGATTTGGATCCTGCTCCTAATCCACCTCA 4139
1173 tPheGluLeuSerValProPheArgGlnGlnHisTyrLeuAlaGlyLeuV 1190
4140 TTTTCCATCGCATGTGATTAAAGCAACATTTGCCCTATATCAGCAATTGTC 4189
1190 aIleuThrGluLeuAlaValIleLeuAsp.....ProAsp 1201
4190 ATAAACCAAGTTAAAGCATTTTAGAAATCTTCCAAAGCCCTGAT 4239
1202 AlaGluGlyLeuPheGlyLeuHisLysLysValIleAsnMetValHisAs 1218
4240 TCC.....TATCAGAAAT..... 4254
1218 nLeuLeuSerSerHisAspSerAspProArgTyrSerAspProGlnIleL 1235
4255 .CTTCTTGCCATATGTGAGCAAGCAGCTGAACAATAATGTTTATAAGA 4303
1235 ysaAlaArgValAlaMetLeuTyrLeuProLeuIleGlyIleIleMetGlu 1251
4304 AGCACAGAAATCTTAAATATATACCTGTGTGTAGTTTATTACTGAAA 4353
1252 ThrValProGln.....LeuTyrAsp.. 1258
4354 GATTATAAAAGTGGCTTAGAGAGAGCTTGGGCTTTGTTCTTCGAGAGCT 4403
1259PheThrGluThrHisAsnGlnArgGlyArgProIleCysIleAlaT 1274
4404 TATTATACTTTGATTCACATATATCAACCAAGGCTTCTGTATCATGG 4453
1274 hrAspAspTyrGluSerGluSer.....GlySerMetIle 1285
4454 ATGTGTCATTACGTAGCTTCTCCCTTGTGTGACTTATTAAGTACAGCTT 4503
1286 SerGlnThrValAlaMet..... 1291
4504 TGGCAGACAGCCGTACTTACTGTAAAGATGCTCTAGAAAAACCATCTTCA 4553
1292 .AlaIleAlaGlyThrSerValProGlnLeuThrArgProGlySerPheL 1308
4554 TGTATTTGTTGGTACACTTATACCCCTTGTGTATGAGCAGGTGAGGTTTC 4603
1308 euleuThrSerThrSerGlyArgGlnHisThrThrPheSerAlaGluSer 1324
4604 AGAAACAGGTATTGACTTGTGAATACTTAGTAGATACAAAGGAT 4653

1325 SerArgSerLeuLeuIleCysLeuLeuTrpValLeuLysAsnAlaAspGl 1341
4654 AATGAAACCTCTATATACAGATTAAAGCTTTTAGATCCCTTTCTCGACCA 4703
1341 uThrValLeuGlnLysTrpPheThrAspLeuSerValLeuGlnLeuAsn. 1357
4704 TGTGTGT.....TTTAAGGATTTGCGTATTACTCAGCAAAAAA 4741
1358ArgLeuLeuAspLeuLeu.....Tyr 1364
4742 TCAATACAGTAGAGAGACCCTTTTCACTCTTGAGAGAAATTAACCATTTT 4791
1365 LeuCysValSerCysPheGluTyrLysGlyLysLysValPheGluArgMe 1381
4792 CTCTCAGTAGT.....GTTTATGATGCACCT 4817
1381 tAsnSerLeuThrPheLysSerLysAspMetArgAlaLysLeuGlu. 1397
4818 TCCATTGACAAGACTTGAAGGACTAAAGGATCTTCGAAGACAACCTGGAAC 4867
1398GluAlaIleLeuGlySerIleGlyAlaArgGlnGlu..... 1409
4868 TACATTAAGATCAGATGGTGGACATTATGAGAGCTTCTCAGCATTAATCCG 4917
1410MetValArgArgSerArgGlyGlnLeuGluArgSerPr 1422
4918 CAAGATGGGATTATGGTGAACACTAGTTGTCAATTGTTGTCAGTTATCCAA 4967
1422 oSer.....Glys 1425
4968 GATGGCAATAAACACACACTGGTGAAAGAAAGATTCTAGAGGCTGTGGAA 5017
1425 eAlaPheGlySerGlnGluAsnLeuArgTrpArgLysAspMetThrHis 1441
5018 GCTGCTTGGA..... 5028
1442 TrpArgGlnAsnThrGluLysLeuAspLysSerArgAlaGluIleGluHi 1458
5029GAAGTGGGTCCGATAGATTCTCTACCATAGCTATACACACA 5069
1458 sGluAlaLeuIleAspGlyAsnLeuAlaThrGluAlaAsnLeu..... 1472
5070 TAGTAAA.....GATGCATCTTATACCAAGGCCCTTAAGTTATTGAAG 5113
1473IleIleLeuAspThrLeuGlu..... 1479
5114 ATAAAGAACTTCAGTGGACCTTCATAATGCTGACCTTACCTGAATAACACA 5163
1480 IleValValGlnThrValSerValThrGluSerLysGluSerIleLeuGl 1496
5164 CTGGTAGAGATTGTGTCAAGTTCGATCAGACAGCTGTTAACCTGTTTGA 5213
1496 yGlyValLeuLysValLeu.....LeuHisSerM 1506
5214 AAACATTTTAGCCACAAAGACTGGACATAGTTTCTGGGAGATTATAGA 5263
1506 eAlaCysAsnGlnSerAlaValTyrLeuGlnHisCysPheAlaThrGln 1522
5264 TGACAACAGATCCCATGCTGGCTATCTACAGCCT..TTTGAACATCA 5310
1523 ArgAlaLeuValSerLysPheProGluLeuLeuPheGluGluGluThrGl 1539
5311 AGAAAAAGTTTTTAGAAGTACCAGATTTGACAAGAAAAACCCCTTTTGA 5360
1539 uGlnCysAlaAspLeuCysLeuArgLeuLeuArgHisCysSerSerSerI 1556
5361 AGGCTGGATGATATA..... 5376
1556 leGlyThrIleArgSerHisProSerAlaSerLeuTyrLeuLeuMetArg 1572
5377AATCTGTGATTCCTCTAAGT 5397

1573 GlnAsnPhgLeuIleGly.....AsnAsnPh 1581
::||| ::||| :::
5398 GAAAAATCATGACATTTGGATAAAGACACTGCTTGTGCTTTTGGACAG 5447
1581 eAlaArgValIysMetGlnValPrometSerLeuSerSerLeuValGlyT 1598
:: ::||| :::: :: |||:: ::
5448 TGGAGGCACAAATGTGAATTCCTCAATTATTAAAGCCCAATGTGTGAAG 5497
1598 hrSerGlnAsnPhgAsnGluGluPheLeuArgSerLeuIysThrIle 1614
::: ::||| ::|||:::
5498 TGAAAACTGACTTT.....TGTACAGACTGTA 5523
1615 LeuThrTyraIaGluGluAspLeuGluLeuArgGluThrThr..... 1628
||| ||| ::|||:: |||:::||||:
5524 CTTCCATATCTGATTCATGATATTCTTACTCCAAGATACAAATGAATCATG 5573
1629PheProAspGlnValGlnAspLeuValPheAsnLeuHisM 1642
::: ::|||:: :::: :::
5574 GAGAAATGTGCTTTCTTACACATGTTCAGGGATTTTTCACCAGCTGTCTTC 5623
1642 etIleLeuSerAspThrValIysMet..... 1650
::|||::||| :::
5624 GACACTTCTGCGCAACGAGCCGATCCACAACCCCTGCAACCTTGATTTCA 5673
1651LysGluHisGlnGluAs 1656
|||:: |||:::
5674 GAGTCAGAGCACTTTTCCGATGCTGTGGATTAATAAATCACAAGAAGAC 5723
1656 pProGluMetLeuIleAspLeuMetTyraArgIleAlaIysGlyTyrglnT 1673
:::||||| ||| |||
5724 AATGCTTGCTGTGTGGACTACATGAGAAGA.....CAA 5758
1673 hrSerProGluArgLeuThr.....TrpLeuGlnAsn 1683
|||:: ||| |||
5759 AGAGACCTTCTTCAAGACAATTTTAATGATGCTTCTGCTGGAT... 5805
1684 MetAlaGlyLysHisSerGluArgSerAsnHisAlaGluAlaIaGln.. 1699
|||:: |||:::||||:
5806TTAAATTATCTAGAAGTTGCCAAGT 5831
1700CysLeuValHisSerAlaAlaLeuValAlaGluTyrlus 1713
||| ::||| ::||| |||:
5832 AGCTCAGTCTTGTGCTGCTCAGCTTACAGCTTTACTC.....TATGCAG 5875
1713 erMetLeuGluAspArgLysTyrlEuPro..... 1722
::: |||::||| :::
5876 AAATCTATGCAGATAAGAAAGATGATGATCAAGAGAAAGAAAGTCTT 5925
1723ValGlyCysValThrPheGlnAsnIleSerSerAs 1734
::|||:::||||:
5926 GCATTTGAAGAGGAAGCCAGAGTACACTATTCTTAGCTTGAGTGAATA 5975
1734 nValLeuGluGluSerAlaValSer..... 1742
: |||||:::||||
5976 AAGTAAAGAGAAACTGGAATAGTTTACAGGATCTTCTTAGAAATCT 6025
1743 ..AspAspValValSerProAspGluGluGlyIleCysSerGly..... 1756
::: ||| |||:
6026 ACAGAAGTATAGGGAGCCAGATAGTTGTATGGCTGTGTGGAGGGAAG 6075
1757LysTyrrPheThrGluSerGlyLeuVa 1765
::: :: |||:
6076 ATGTACAACCCATTACTAGACTAGCAACATATGAACACAGCAATGTG 6125
1765 IglyLeu.....LeuGluGlnAla..AlaAlaSerP 1775
||| ||||| ||| ::|||
6126 GGGCAAAAGCCCTAGTAACATATGACCTCGAAACAGCAATCCCTCATCAA 6175
1775 heSerMetAlaGlyMetTyrgluAlaValAsn..... 1785
::||| |||:::
6176 CACGCCAGGCGGAATCATTCAGGCTTGCAGAATTTGGACTCTGCCAT 6225
1786GluValTyrlLysValLeuIleProIleHisGluAlaAsnArgAs 1800

6226 ATTCTTTCCTCTAT.....TTAAAGGATTTGGATTATGAAATAAAGA 6269
::||| ||| :::
1800 p..AlaLysLysLeuSerThrIleHisGlyLysLeuGlnGluAlaPhe. 1815
| :::: ||| |||
6270 CTGTGTCTGTAAGTGAAGAACTTCAT.....TACCAAGCAGCATGGA 6313
1816SerLysIleValHisGln 1821
||| ||| |||:
6314 GGAATATGACGTGGACCATTTGCACCTTCGTCAGCAAGAAGATGAAGA 6363
1822 SerThrGlyTrpGluArgMetPheGlyThrTyrr..... 1832
::: ||| ::::: |||
6364 ACCAGTTACCATGATCATTTGTACAATGCTCTACAATCTCTAAGACAG 6413
1833 ...PheArgValGlyPheTyrglyThrLysPheGlyAspLeuAspGluG 1848
||| :::: |||:
6414 AGAATTCCTACATTTTATGAAGTCTCAAAATATGCCAGAGATTAAGAAG 6463
1848 InGluPheValTyrlLysGlu.....Pro 1855
||| :: |||:
6464 TGAAGAGATGTGTAAAGCGCAGCCCTTGAGTGTGTATTCGCTCATCCC 6513
1856 AlaIleThrLysLeuAlaGluIleSerHisArgLeuGluGlyPheTyrgl 1872
::: |||:: |||:
6514 ACACCTTAGCAGTTGCAGGCCATTGGA...GAGCTGGAAGCAATT...Gg 6557
1872 yGluArgPheGlyGluAspValVal..... 1880
||| |||::: |||:
6558 GGAGCTTTTCTCAAGATCACTCACACATAGACAACCTCTGAAAGTATATA 6607
1881GluValIleLysAspSer..... 1886
::: ||| ||| |||
6608 TTAAGTGCAGAAACACTCCACAGCTTCTCAAGACAGATGATTTAGTTT 6657
1886 1886
6658 CAGAGCCTATCATGCGCTCTACGCACACATGATTTTGGAGATCCTGATGA 6707
1887AsnProValAspLysCysLysLeuAsp..ProAsnL 1898
||| ||| ::||| |||
6708 AAAGGAATGACAACTCACAAAGAGAATGTATTAAAGACATTTCCACCA 6757
1898 ysaIaTyrlIleGlnIleThrTyrrValGluProTyrrPheAspThrTyrglu 1914
|| :::: |||:
6758 AACACCTTGTAGACTCTCTATACTGCCAGAACTTTCAAGAACAACCTCAG 6807
1915 MetLysAspArgIleThrTyrrPheAspLysAsnTyrrAsn 1927
::: ||| :: ||| |||
6808 CTCCTGAAGGCAATATTTCAAATTAACAGTACAAT 6846
seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-952-127-2
seq_documentation_block:
; Sequence 2, Application US/08952127
; Patent No. 6211336
; GENERAL INFORMATION:
; APPLICANT: Shiloh, Yosef
; APPLICANT: Tagle, Danilo A.
; APPLICANT: Collins, Francis S.
; TITLE OF INVENTION: ATAXIA-TELANGIECTASIA GENE
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kohn & Associates
; STREET: 30500 No. 6211336thwestern Hwy., Suite 410
; CITY: Farmington Hills
; STATE: Michigan
; COUNTRY: U.S.
; ZIP: 48334
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

1173 tphegluLeuSerValPropheargInglInHisTyrLeuAlaGlyLeu 1190
|||
4140 TTTTCCATCGCATGTGATTAAGCACACATTTGCCATATATCAGCAATTGTC 4189
|||
1190 alleuthrGluLeuAlaValIleuAsp.....ProAsp 1201
|||
4190 ATAAACCAAGTTAAAAAGCATTTTAGAAATCTTTCCAAAAAGCCCTGAT 4239
|||
1202 AlaGlyLeuPheGlyLeuHisLysValIleAsnMetValHis 1218
:::
4240 TCC.....TATCAGAAATTT..... 4254
1218 nLeuLeuSerSerHisAspSerAspProArgTyrSerAspProGlnIle 1235
|||
4255 .CTCTGCCATATGTGAGCAGACGCTGAACAATATATGTTATAAGA 4303
|||
1235 ysAlaArgValAlaMetLeuTyrLeuProLeuIleGlyIleMetGlu 1251
|||
4304 AGCACAGAAATCTTAAATATATCACCTGTTGTTAGTTATTAAGTGA 4353
|||
1252 ThrValProGln.....LeuTyrAsp.. 1258
:::
4354 GATATAAAAAGTGGCTTAGAGAGAGCTTGGCCTTGTGTTCTTCGAGACGT 4403
|||
1259 ...PheThrGluThrHisAsnGlnArgGlyArgProIleCysIleAla 1274
:::
4404 TATTTATACTTGTGATTCATATATCAACAAAGCCTTCTGTATCATGG 4453
1274 hrAspAspTyrGluSerGluSer.....GlySerMetIle 1285
:::
4454 ATGTGCATTTAGCTGCTCTCCCTTGTGTGACTTATTAAGTCAGGTT 4503
1286 SerGlnThrValAlaMet..... 1291
|||
4504 TGCCAGACAGCCGTGACTTACTGTAGAGATGCTCTAGAAACCATCTTCA 4553
1292 .AlaIleAlaGlyThrSerValProGlnLeuThrArgProGlySerPhe 1308
:::
4554 TGTATTTGTTGTTACACTTATACCCCTTGTGTATGAGCAGGTGAGGTTTC 4603
1308 euLeuThrSerThrSerGlyArgGlnHisThrPheSerAlaGluSer 1324
:::
4604 AGAAACAGGTATTTGACTTGTGAAATACTTAGTATAGATAACAAGGAT 4653
1325 SerArgSerLeuLeuIleCysLeuLeuTyrValLeuLysAsnAlaAspG 1341
:::
4654 AATGAAACCTTATATACAGATTAAGCTTTTAGATCCTTTTCCTGACCA 4703
1341 uThrValLeuGlnLysTrpPheThrAspLeuSerValLeuGlnLeuAsn. 1357
:::
4704 TGTGTGT.....TTTAAAGATTTGCGTATTACTCAGCAAAAAA 4741
1358ArgLeuLeuAspLeuLeu.....Tyr 1364
4742 TCAATATACAGTAGAGACCCTTTTCACTCTTGAGAGAAATTAACCATTTT 4791
1365 LeuCysValSerCysPheGluTyrLysGlyLysValPheGluArgMe 1381
|||
4792 CTCTCAGTAAGT.....GTTTATGATGCACAT 4817
1381 tAsnSerLeuThrPheLysLysSerLysAspMetArgAlaLysLeuGlu. 1397
:
4818 TCCATTGACAAAGACTTGAAGACTTAAAGATCTTCGAAGACAACTGGAAC 4867
1398GluAlaIleLeuGlySerIleGlyAlaArgGlnGlu..... 1409
4868 TACATAAAGATCAGATGTGAGACATTAAGAGCTTCTCAGGATTAATCCG 4917
1410MetValArgArgSerArgGlyGlnLeuGluArgSerPr 1422
4918 CAAGATGGGATTTATGTTGAAGACTAGTTGTCAATTTGTTGCAGTTATCCAA 4967
1422 oSer.....Glys 1425

4968 GATGCAATAAACACACACTGGTGA AAAAGAGTTCTAGAGGCTGTGGAA |||
1425 erAlaPheGlySerGlnGluAsnLeuArgTyrArgLysAspMetThrHis 1441
|||
5018 GCTGCTGGGA..... 5028
1442 TrpArgGlnAsnThrGluLysLeuAspLysSerArgAlaGluIleGluH 1458
:::
5029GAAGTGGCTCCTATAGATTTCTCTACCATAGCTATACACA 5069
1458 sGluAlaLeuIleAspGlyAsnLeuAlaThrGluAlaAsnLeu..... 1472
5070 TAGTAAA.....GATGCATCTTATACCAAGGCCCTTAAGTTATTTGAAG 5113
1473IleIleLeuAspThrLeuGlu..... 1479
5114 ATAAAGAACTTCAGTGGACCTTTCATATAGCTGACCTGAATACACA 5163
1480 IleValValGlnThrValSerValThrGluSerLysGluSerIleLeuG 1496
5164 CTGGTAGAAGATTTGTCAAAAGTTCGATCAGCAGCTGTACCSTGTTGAA 5213
1496 yGlyValLeuLysValLeu.....LeuHisSerM 1506
5214 AAACATTTTAGCCACAAAGACTGGACATAGTTCTGGGAGATTTATAGA 5263
1506 etAlaCysAsnGlnSerAlaValTyrLeuGlnHisCysPheAlaThrGln 1522
5264 TGACAACAGATCCATGCTGGCCTATCTACAGCCT..TTAGAACAATCA 5310
1523 ArgAlaLeuValSerLysPheProGluLeuLeuPheGluGluThrG 1539
5311 AGAAAAAGTTTPTAAGAGTACCAGATTTGACAAAGAAACCCCTTTTGA 5360
1539 uGlnCysAlaAspLeuCysLeuArgLeuLeuArgHisCysSerSerSer 1556
5361 AGCCCTGATGATATA..... 5376
1556 IeGlyThrIleArgSerHisProSerAlaSerLeuTyrLeuLeuMetArg 1572
5377AATCTGTGATTTCCCTAGAT 5397
1573 GlnAsnPheGluIleGly.....AsnAsnPh 1581
5398 GAAATCATGACATTTGGATTAAGACACTGACTTGTGCTTTTGGACAG 5447
1581 eAlaArgValLysMetGlnValPrometSerLeuSerSerLeuValGlyT 1598
5448 TGGAGGCACAAAGATGTAATTTCTCAATTAATTAAGCCAAATGTGTGAAG 5497
1598 hrSerGlnAsnPheAsnGluGluPheLeuArgArgSerLeuLysThrIle 1614
5498 TGA AAAACTGACTTT.....TGT CAGACTGTA 5523
1615 LeuThrTyrAlaGluGluAspLeuGluLeuArgGluThr..... 1628
5524 CTTCATACTTGTGATTCATGATATTTTACTCCAGATACAAATGAATCATG 5573
1629PheProAspGlnValGlnAspLeuValPheAsnLeuHisM 1642
5574 GAGAAATCTGCTTTCTACACATGTT CAGGGATTTTTCACCAAGCTGTCTTC 5623
1642 etIleLeuSerAspThrValLysMet..... 1650
5624 GACACTTCTCGCAAAAGCAGCCGATTCACAACCCCTGCCAAACTTGATTC 5673
1651LysGluHisGlnGluAs 1656
5674 GAGTCAGAGCACTTTTTCGATGCTGTTTGGATTA AAAAATCACA AAAGAAC 5723
1656 pProGluMetLeuIleAspLeuMetTyrArgIleAlaLysGlyTyrGlnT 1673

```

5724 AATGCTTGGCTGTGTGTGGACTACATGAGAAGA.....CAAA 5758
1673 hrSerProGluArgLeuThr.....TrpLeuGlnAsn 1683
      |||::: |||
5759 AGAGACCTTCTTCAGGAACAATTTTAAATGATGCTTTCTGGCTGGAT... 5805
1684 MetaGluGlyHisSerGluArgSerAsnHisAlaGluAlaGln.. 1699
5806 .....TTAAATTAATCTAGAAGTTGCCAAGGT 5831
1700 .....CysLeuValHisSerAlaAlaLeuValAlaGluTyrLeus 1713
      ||| ::||| ::|||::: |||
5832 AGCTCAGTCTTGTGCTGCTCAGCTTACAGCTTTACTC.....TATGCAG 5875
1713 erMetLeuGluAspArgLysTyrLeuPro..... 1722
      :::: |||::||| :::
5876 AAATCTATGCAGATAAGAAAGATGATGATCAAGAGAAAGAAGTCTT 5925
1723 .....ValGlyCysValThrPheGlnAsnIleSerSerAs 1734
      ::|||::: |||::: |||::: |||:::
5926 GCATTGGAAGAGGAAGCCAGAGTACAACTATTCTAGCTTGAGTGAAAA 5975
1734 nValLeuGluGluSerAlaValSer..... 1742
      : |||||::: |||
5976 AAGTAAGAAGAAACTGGAATAGTTTACAGGATCTTCTTAGAAATCT 6025
1743 ..AspAspValValSerProAspGluGluGlyIleCysSergly..... 1756
      ::::: ::|||::: |||::|||
6026 ACAGAAGTATAGGGGAGCCAGATAGTTTGTATGGCTGTGTGGAGGGAAG 6075
1757 .....LysTyrPheThrGluSerGlyLeuVal 1765
6076 ATGTTTACAACCCATTACTAGACTACGAAACATATGAAACAGAAAGCAATGTG 6125
1765 IgGlyLeu.....LeuGluGlnAla..AlaIleSerP 1775
      ||| ||||| ||| ::|||
6126 GGGCAAAAGCCCTAGTAACATATGACCTCGAAACAGCAATCCCTCATCAA 6175
1775 heserMetaIleGlyMetTyrGlnAlaValAsn..... 1785
      ::|||::: ::|||:::
6176 CAGCCAGGAGGAATCATTCAGGCTTGCAGAATTTGGAGCTTGCCAT 6225
1786 .....GluValTyrLysValLeuIleProIleHisGluAlaAsnArgAs 1800
      ::||| ||| ::: |||::|||
6226 ATTCCTTCCCTCTAT.....TTAAAGGATTGGATTATGAATAAAGA 6269
1800 P...AlaLysLysLeuSerThrIleHisGlyLysLeuGlnGluAlaPhe. 1815
      | ::: ::|||::: ::||| ||| |||:::
6270 CTGGTGTCTCTGAAGCTAGAGAAGACTTCAT.....TACCAGCAGCATGGA 6313
1816 .....SerLysIleValHisGln 1821
6314 GGAATATGCAGTGGGACCATTCGACTCCGTCAGCAAGAAGATGAGAAGA 6363
1822 SerThrGlyTyrPgluArgMetPheGlyThrTyr..... 1832
      ::::: ||| ::::: |||
6364 ACCAGTTACCATGAATCATTTGTACAATGCTCTACAATCTCTAAGAGACAG 6413
1833 ...PheArgValGlyPheTyrGlyThrLysPheGlyAspLeuAspGluG 1848
      ||| ::: ::: |||::: |||
6414 AGAATTCCTCTACATTTTATGAAGTCTCAAAATATGCCAGAGTAAGAAG 6463
1848 InGluPheValTyrLysGlu.....Pro 1855
      ||| ::: |||::: |||
6464 TGGAAAGAGATGTGTAAAGCGCAGCCTTGAGTCTGTGATTCGCTCTATCCC 6513
1856 AlaIleThrLysLeuAlaGluIleSerHisArgLeuGluGlyPheTyrGln 1872
      ::::: ||| |||::: ::||| |||::: |||
6514 ACACTTAGCAGGTTCAGGCCCATTTGGA..GAGCTGGAAGCATT...GG 6557
1872 yGluArgPheGlyGlyLeuAspValVal..... 1880
      |||| |||::: |||::: |||:::
6558 GGAAGCTTTTCTCAAGATCAGTCAACATAGACAACACTCTGTGAATATATA 6607

```

```

1881 .....GluValIleIleLysAspSer..... 1886
6608 TTAAGTGGCAGAAACACTCCCGAGCTCTCAAGGACAGTGATTTTACTTTT 6657
1886 ..... 1886
6658 CAGAGCCTATCATGGCTCTACGCACAGTCATTTTGGAGATCCTGATGGA 6707
1887 .....AsnProValAspLysCysLysLeuAsp..ProAsnL 1898
6708 AAAGGAATGGACAACACTCACCAAGAAGATGATTTAAGGACATTTCTCACCA 6757
1898 ysaIatYrIleGlnIleThrTyrValGluProTyrPheAspThrTyrGlu 1914
      ||      :::::      ::      ||      ::      ::
6758 AACACCTTGTAGAACCTCTCTATACTGGCCAGAAGCTTTCAGAAGACACTCAG 6807
1915 MetLysAspArgIleThrTyrPheAspLysAsnTyrAsn 1927
      ::      ::::||      ::      ||:::||::||
6808 CTCCCTGAAGGGCAATATTTCAATTTAAACAGTACAAT 6846

```

```
seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-952-014C-2
```

seq_documentation_block:

Patent No. 6265158

GENERAL INFORMATION

; APPLICANT: Shiloh, yosef
 ; TITLE OF INVENTION: ATAXIA-TELANGIECTASIA GENE AND ITS

; TITLE OF INVENTION: GENOMIC ORGANIZATION

CORRESPONDENCE ADDRESS:

STREET: 30500 No. 6265158thwestern Hwy., Suite 410

CITY: Farmington Hills

COUNTRY: U.S.

ZIP: 48334

MEDIUM TYPE: Floppy

COMPUTER: IBM PC compatible

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/952,014C

FILING DATE: ;

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Kohn, Kenneth I.

REGISTRATION NUMBER: 30,995

REFERENCE/DOCKET NUMBER: 2290.00028;

TELECOMMUNICATION INFORMATION: 810 530 5050
; EDITED BY: 810 530 5050

TELEPHONE: 810-539-5050
TELEFAX: 810-530-0066

TELEFAX: 810-5339-5055

```

; INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:

```

SEQUENCE CHARACTERISTICS: 9171 base pairs

LENGTH: 91/1 base pair
TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

ORIGINAL SOURCE:

ORGANISM: Homo sapiens

POSITION IN GENOME:

CHROMOSOME/SEGMENT: 11q22-23

US-08-952-014C-2

alignment_scores:

Quality:	150.50	Length:	2063
Ratio:	0.175	Gaps:	99

Percent Similarity: 41.638 Percent Identity: 17.402

alignment block:
US-09-737-246-2 x US-08-952-014C-2 ..
Align seg 1/1 to: US-08-952-014C-2 from: 1 to: 9171

267 ValIysCysLeuSerLeuYspheGluIleGluProIlePheAl 283
||| |||||::||| ::::||| ||| :::
1615 GTATCGCTGTTGACTTTGGCACTGACCACCACTATAGTTCAGGAACGGT 1664
283 aSerLeuAlaLeuTyraSpValIysGluYsLysIleSerGluAsnP 300
:::~::~~::~ |||::~::~ |||
1665 AAAAATGGGAATA.....GAGCAAAATATGTGTAAGTAATA 1702
300 heTyRpheAspleuAsnSerGluGlnMetIysGlyLeu..... 313
|||::|||::~::~~::~ ||||| |||||
1703 GAAGCTTTCTTTAAAGCAATCAATAATGAATGGCTTATCTATCAG 1752
314ArgProHisValProProAlaAlaIleTh 323
1753 TTAGAGGGTGACTTAGAAATAGCACAGAAAGTGCCTCCA..... 1791
323 rThrLeuAlaArgSerAlaIlePheSerIleThrTyProSerGlnAspV 340
|||::~::~~::~ |||::~::~~::~ |||
1792ATTCTTCACAGTAATTTCCTCAT..... 1815
340 alPheLeuValIleLysLeuGluYsValLeuGlnGlnIlyaspIleGly 356
::~::~~::~ |||||::~::~~::~ ||| :::
1816CTTGACTGGAGAAATCTTGTGAGTCTCAGCTATGAAA 1854
357 GluCysAlaGluProTyrMetIlePheLysGluAlaAspAlaThrLysAs 373
::~::~~::~ |||::~::~~::~ |||::~::~~::~ |||
1855 AACTGTAAGCTGCAATGAATTTTCCAAAGCGTGCAGAAATGTGAACA 1904
373 n.....LysGluLysLeuGluYsLeuYsSerGlnAlaAspGlnPheC 388
: |||::~::~~::~ ||| |||::~::~~::~ |||
1905 CCACCAAAAGATTAAGAAGAACTTTCATCTCAGAAAGTAGAAGAACTAT 1954
388 ysgLnaArgLeuGlyLysTyraArgMetProPheAlaTrpAlaIleHis 404
::~::~~::~ ||| |||
1955 TTCTTCAGACAACTTTTGACACAGATGACTTT..... 1986
405 LeuMetAsnIleValSerSerAlaGlySerLeuGluArgAspSerThrGl 421
::~::~~::~ ||||| ::~::~~::~ |||::~::~~::~ |||
1987 ..TTAACCATTTGTGAGAGAATGTGT..ATAGAAAAGCACCACTCCAG 2030
421 uValGluIleSerThrGlyGluArgLysGlySerTrpSerGluArgArgA 438
::~::~~::~ |||::~::~~::~ |||
2031 TATGGCTTCTCTGTCACACAGAACTCTC..... 2058
438 snSerSerIleValGlyArgArgSerLeuGluArgThrThrSerGlyAsp 454
::~::~~::~ |||
2059AAGGAATCACTGGATCGCTGTCTGTGGGA... 2088
455 AspAlaCysAsnLeuThrSerPheArgProAlaThrLeuThrValThrAs 471
2088 2088
471 nphePheLysGlnGluGlyAspArgLeuSerAspGluAspLeuTyrlYsp 488
||| |||||::~::~~::~ ||| ::~::~~::~ |||
2089TTATCAGAACAGCTTCTGAATAATT 2113
488 heLeuAlaAspMetArgArgProSerSerValLeuArgArgLeuArgPro 504
::~::~~::~ ||| ::~::~~::~ |||
2114 ACTCATCTGAGATTACAATTCAGAAACTCTGTCCGGGTGTTCACGTCCTT 2163
505 IleThrAlaGlnLeuLysIleAspIleSerProAlaProGluAsnProHi 521
::~::~~::~ |||
2164 TTGGTGGGTGTCTTGGCTGC..... 2184
521 sTyrcysLeuThrProGluLeuGlnValLysLeuTyrlProAspSera 538
||| ||| ::~::~~::~ |||
2185 .TACTGTTACATGGGTGTAATAGCTGAAGAGGAAGCATATTAAGTCAGAAAT 2233

538 rgValArgProThrArgGluIleLeuGluPheProAlaArgspValTy 554
::~::~~::~ |||
2234 TATTCAGAAAGCCAAAGTCTCTAATGCAATGTGCAGAGAAAGTATC... 2280
555 ValProAsnThrThrTyraArgAsnLeuLeuTyrlIeTyrlProGlnSerIe 571
2281ACTCTGTTAAAAAT.....AAGAC 2300
571 uAsnPheAlaAsnArgGlnGlySerAlaArgAsnIleThrValLysValG 588
||| ||| ||||| |||||::~::~~::~ |||
2301 AAATGAGGAATTCAGAAATTGCTTCCCTTGAGAAATATG.....ATGC 2341
588 lnPheMetTyrlGlyGluAspProSerAsnAlaMetProValIlePheGly 604
|||::~::~~::~ |||
2342 AGCTATGT.....ACACGTTGCTTG..... 2361
605 LysSerSerCysSerGluPheSerLysGluAlaTyrlThrAlaValValTy 621
|||::~::~~::~ |||
2362 ..AGCAACTGTACCAAG..... 2376
621 rHisAsnArgSerProAsp.....PheHisGluGluI 632
::~::~~::~ |||
2377AAGAGTCCAAATAAGATGCACTGCGCTTTTTCCTCGATTGT 2419
632 leLysValLysLeuProAlaThrLeuThrAspHisHisIleLeuPhe 648
::~::~~::~ ||||| ::~::~~::~ ||| |||
2420 TAACATCAAAAGCTAATGAATGACATTCAGATATTTGTAAGAGTTTAGCA 2469
649 ThrPheTyrlHisValSerCysGlnGlnLysGlnAsnThrProLeuGluTh 665
::~::~~::~ ||| ::~::~~::~ |||
2470 TCCTTCATCAAAAGCCATTTTGACCGGTGAGAAAGTAGAATCAATGAAGA 2519
665 rProValGlyTyrlThrTrpIleProMetLeuGlnAsnGlyArgLeuLys. 681
::~::~~::~ ||| |||
2520 TGATACT.....AATGGAATCTTAATGG 2542
682ThrGlnPheCysLeuProValSerLeuGluYsProGln 696
||| ||| ::~::~~::~ |||
2543 AGGTGAGAGATCATCATGAAATCTAATTAAAGATTAACGATTACCTGATAGT 2592
697 AlaTyrlSerValLeuSerProGluValProLeuProGlyMetLysTrpVa 713
::~::~~::~ ||| ::~::~~::~ |||
2593 AGTGTTAGT.....GATGCAAAAGAACTTGA..... 2619
713 lAspAsnHisLysGlyValPheAsnValGluValAlaAlaValSerSerI 730
::~::~~::~ ||| ::~::~~::~ |||
2620 .GAGAGCCAAAGTACCATAGGTGCCATTAATCCTTTAGCTGAAGAAATATC 2668
730 leHisThrGlnAsp...ProTyrlLeuAsp.....LysPhePheAlaLeu 743
::~::~~::~ ||||| ::~::~~::~ ||| |||||::~::~~::~ |||
2669 TGTCAAAAGCAGATCTACTTTTCTTAGACATGCTCAAGTTCCTGTGTTTG 2718
744 ...ValAsnAlaLeuAspGluHisLeuPheProValArgIleGlyAspMe 759
|||::~::~~::~ ||| ||| ::~::~~::~ |||
2719 TGTGTAACTACTGCTCAGACCAATACTGTGTCTTTAGGGCAGCTGATAT 2768
759 tarGileMetGluAsnAsnLeuGluAsnGluLeuLysSerSerIleSera 776
||| ||| ::~::~~::~ |||
2769 TCGG.....AGGAAATGTTAATGT 2788
776 lAleuAsnSerSerGlnLeuGluProValValArgPheLeuHisLeu... 791
::~::~~::~ ||||| ||||| ||||| |||||
2789 TAAATTGATTCTAGACGCTAGAACCT..ACCAAAATCCCTCCACCTGAT 2835
792LeuLeuAspLysLeu..... 796
2836 ATGTATCTAATGCTTTAAAGAGAGCTTCTGAGAGAAGATACCCCTTGCC 2885
797IleLeuLeuValIleArgProProValIleAlaGlyGlnI 810
::~::~~::~ ||| ::~::~~::~ |||
2886 AATGGAAGATGTTCTTGAACCTCTGAACCA.....CTATCCAAATG 2926
810 leValAsnLeuGlyGlnAlaSerPheGluAlaMetAlaSerIleIleAsn 826

```
2927  TGTGTTCTTTGTATCGTCGTGACCAAGATGTTGTAAACTATTTTAAC 2976
      ::  ::|||  ::  ::  ::|||  ::|||  ::|||
827   Arg...LeuHis.....LysAsnLeu..GluGlyAsnHisAspGlnH 839
      ::  |||||  |||||  |||||  |||||  |||||  |||||
2977  CATGTCCTTCATGTAGTAAAAACCTAGGTCAACCAATATGACCTCTGA 3026
      ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
839   sGlyArgAsnSerLeuLeuAlaSerTyrIleHisTyrValPheArgLeuP 856
      ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
3027  GAACACAAGGATGCTCAAGACAGTTCCTTACAGTAATT..... 3066
      ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
856   roAsnThrTyrProAsnSerSerSerProGlyProGlyGlyLeuGly 872
      ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
3067  .....GGA 3069
      ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
873   SerValHisTyrAlaThrMetAlaArgSer.....AlaValArgPr 886
      ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
3070  GCATTTTGGCATCTAACAAAGAGAGAAATATATATTCTCTGTAAAGAT 3119
      ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
886   oAlaSerLeuAsnLeuAsnArgSerArgSerLeuSerAsnSerAsnPro. 902
      |||  ::|||  ::|||  ::|||  ::|||  ::|||
3120  GGCCCTAGTAAATTGC.....CTTAAACTTTGCTTGAGGCTGATCCTT 3163
      |||  ::|||  ::|||  ::|||  ::|||  ::|||
903   .....AspIleSerGlyThrProThrSerProAsp 912
      ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
3164  ATTCAAATGGGCCATTCTTATGTATGGGA...AAAGACTTTCCTGTA 3210
      ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
913   AspGluValArgSer.....Ile 918
      ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
3211  AATGAAGTATTTACACAATTCTTGCTGACATCATCACCAGTTCGCAT 3260
      ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
918   eIleGlySerLysGlyLeuAspArg..... 926
      ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
3261  GTTGGCTGCAGAGTCATCAATAGATTGTTCCAGGACACGAGGAGATP 3310
      ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
926   ..... 926
      ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
3311  CTTCCAGGTTACTGAAGCACTTCCTTTGAAGCTTCAGCAAAACAGCTTTP 3360
      ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
927   SerAsnSerTyrValAsnThrGlyGlyProLysAlaAlaProTyrGlySe 943
      ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
3361  GAAAATGCATACTG.....AAAGCTCAGGAAGAAATGAG 3395
      ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
943   rAsnProSerProSerAlaGluSerThrGlnAlaMetAspArgSerCysA 960
      ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
3396  AGAATGTCCCAAGTAGCTGAGAACCCGCAAACTTTGGATGAATTTATA 3445
      ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
960   snArgMetSerSerHisThrGluThrSerSerPheLeuGlnThrLeuThr 976
      |||||  |||||  |||||  |||||  |||||  |||||
3446  ATAGA.....AATCTGTTTACTGACGTTGATA 3474
      |||||  |||||  |||||  |||||  |||||  |||||
977   GLYArgLeuProThrLysLysLeuPheHisGluGluLeuAlaLeuGlnTr 993
      ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
3475  GCTGTGTTTATCTCTGTAGCCCTATCTGCGAAAAACAGGCTTTG...TT 3521
      ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
993   pValValCysSerGlySerValArgGluSerAlaLeuGlnGlnAlaTrp 1010
      ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
3522  TGCCCTGTGTAA...TCTGTGAAGAGAAATGATTAGAACCTTCAC... 3564
      ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
1010  hePhePheGluLeuMetValLysSerMetValHisHisLeuTyrPheAsn 1026
      ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
3565  .....CTGTGAAAAAGGTTTGTAGAGAAAGTT..... 3591
      ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
1027  AspLysLeuGluAlaProArgLysSerArgPheProGluArgPheMetAs 1043
      ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
3592  .....TCTGAACCTTTTGGATATAGACGTTTAGA 3620
      ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
1043  pAspIleAlaAlaLeuValSerThrIleAlaSerAspIleValSerArgP 1060
      ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
3621  AGACTTT.....ATGGCATCTCATTTA..... 3642
      ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
1060  heGlnLysAspThrGluMetValGluArgLeuAsn..... 1071
      |||||  |||||  |||||  |||||  |||||  |||||
```

```
3643  .....GATTATCTGTTTTCGAATGGCTAAATCTCAAGATACTGAA 3684
      ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
1072  .....ThrSerLeuAlaPhePheLeuAsnAspLeuLeuSerValMe 1085
      ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
3685  TACAACCTTATCTCTCTTTCCTTTTATTTTATTAACCTACACAATATTGA 3734
      ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
1085  tAspArgGlyPheValPheSerLeuIleLysSerCysTyrLys..... 1099
      |||||  |||||  |||||  |||||  |||||  |||||
3735  GGAT.....TTCATAGATCTTGTATTAAGGTTTGA 3766
      ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
1099  ..... 1099
      ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
3767  TTCCACATCTGCTGATTAGAGTCATTTTGATGAGGTGAAGTCCATTGCT 3816
      ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
1099  ..... 1099
      ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
3817  AATGAGATTCAAGAGAGACTGGAAGTCTTCTAACAGACTGCTTCCAAA 3866
      ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
1099  ..... 1099
      ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
3867  GATTCTTGTAAATATCTTCCTTATTTTGGCTATGAGGCTACCAGAGACA 3916
      ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
1100  .....GlnValSerSerLysLeuTyrSerLeuPro 1109
      ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
3917  GTGGATGGCACAGCAAGAGAGACTGCTACCAGGCTATGATATGCTT 3966
      ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
1110  AsnProSerValLeuValSerLeuArgLeuAspPheLeuArgIleIleCy 1126
      ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
3967  AAAAGTGAAGAACTTATTTGGGAAAAACAGATTGAT..... 3999
      ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
1126  sSerHisGluHisTyrValThrLeuAsnLeuPro.....CysSerL 1140
      |||||  |||||  |||||  |||||  |||||  |||||
4000  .....CACTTATTCATTAGTAATTATTCACAGAGATTGTGTGGAGT 4039
      ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
1140  euLeuThrProProAlaSerProSerProSerValSerSerAlaThrSer 1156
      |||||  |||||  |||||  |||||  |||||  |||||
4040  TATTGATGACGTTTACATGAGCCAGCAAAATTTCTAGTCCAGTCAAGACT 4089
      ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
1157  GlnSerSerGlyPheSerThrAsnValGlnAspGlnLysIleAlaAsnMe 1173
      ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
4090  GACCTCTGTGACTTTTTCAGGGGATTTGCATCTGCTCCTAATCCACCTCA 4139
      ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
1173  tPheGluLeuSerValProPheArgGlnGlnHisTyrLeuAlaGlyLeuV 1190
      |||||  |||||  |||||  |||||  |||||  |||||
4140  TTTTCATCGCATGTGATTAAGCAACATTTGCCATATATCAGCAATTGTC 4189
      ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
1190  alLeuThrGluLeuAlaValIleLeuAsp.....ProAsp 1201
      |||||  |||||  |||||  |||||  |||||  |||||
4190  ATAAACCAAGTTAAAGCATTTTAGAAATTCCTTCCAAAGCCCTGAT 4239
      ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
1202  AlaGluGlyLeuPheGlyLeuHisLysLysValIleAsnMetValHisAs 1218
      ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
4240  TCC.....TATCAGAAATTT..... 4254
      ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
1218  nLeuLeuSerSerHisAspSerAspProArgTyrSerAspProGlnIleL 1235
      |||||  |||||  |||||  |||||  |||||  |||||
4255  .CTTCTTGCCATATGTGAGCAAGCAGCTGAACAATAATGTTTATTAAGA 4303
      ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
1235  ysAlaArgValAlaMetLeuTyrLeuProLeuIleGlyIleIleMetGlu 1251
      |||||  |||||  |||||  |||||  |||||  |||||
4304  AGCACAGAAATCTTAAATATATACACCTGTTGTAGTTATTTACTGAAA 4353
      ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
1252  ThrValProGln.....LeuTyrAsp.. 1258
      ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
4354  GATATAAAAAGTGCTTAGAGAGACTTGGGCTTTGTTCTTCGAGAGCT 4403
      ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
1259  ....PheThrGluThrHisAsnGlnArgGlyArgProIleCysIleAlaT 1274
      ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
4404  TATTATACTTTGATTTCACTATATACACCAAGGCCCTTCTGTATCATGG 4453
      ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
1274  hrAspAspTyrGluSerGluSer.....GlySerMetIle 1285
      ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
4454  ATGTGTCATTACGTAAGCTTCTCCCTTGTGTGACTTATTAAGTCAGGTT 4503
      ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
```

1286 SerGlnThrValAlaMet..... 1291
1292 .AlaIleAlaGlyThrSerValProGlnLeuThrArgProGlySerPheL 1308
4554 TGTATATGTTGGTACACTTATACCCCTTGTGTATGACGAGGTGAGGTTTC 4603
1308 euleuThrSerThrSerGlyArgGlnHisThrPheSerAlaGluSer 1324
4604 AGAAACAGGATTGTGACCTTGTGAATAGTATGATGATTAACAAGGAT 4653
1325 SerArgSerLeuLeuIleCysLeuLeuTrpValleuLysAsnAlaAspG1 1341
4654 AATGAAAACTCTATATCAGCATTAAGCTTTTGAATCCTTTTCCGTGACCA 4703
1341 uThrValleuGlnLysTrpPheThrAspLeuSerValleuGlnLeuAsn. 1357
4704 TGTGTGTT.....TTTAAGGATTTGGGTATTTACTCAGCAAAAAA 4741
1358ArgLeuLeuAspLeuLeu.....Tyr 1364
4742 TCAATATACAGTAGAGACCCTTTTCACCTCTTGAGGAAATTAACCATTTT 4791
1365 LeuCysValSerCysPheGluTyrLysGlyLysValPheGluArgMe 1381
4792 CTCTCAGTAACT.....GTTTATGATGCACCT 4817
1381 tAsnSerLeuThrPheLysLysSerLysAspMetArgAlaLysLeuGlu. 1397
4818 TCCATTGACAAGACTTGAAGACTAAAGGATCTTCGAAGACAACTGGAAC 4867
1398GluAlaIleLeuGlySerIleGlyAlaArgGlnGlu..... 1409
4868 TACATAAAGATCAGATGGTGACATTATGAGAGCTTCTCAGGATTAATCCG 4917
1410MetValArgArgSerArgGlyGlnLeuGluArgSerPr 1422
4918 CAAGATGGGATTATGGTGAACACTAGTTGTCAATTTGTGCAGTTATCCAA 4967
1422 oSer.....Glys 1425
4968 GATGGCAATAAACACACTGGTGAAAAAGAGTTCTAGAGGCTGTTGGAA 5017
1425 erAlaPheGlySerGlnGluAsnLeuArgTrpArgLysAspMetThrHis 1441
5018 GCTGCTTGGGA..... 5028
1442 TrpArgGlnAsnThrGluLysLeuAspLysSerArgAlaGluIleGluHi 1458
5029GAAGTGGTCCCTATAGATTTCCTCCATAGCTATATACAACA 5069
1458 sGluAlaLeuIleAspGlyAsnLeuAlaThrGluAlaAsnLeu..... 1472
5070 TAGTAAA.....GATGCATCTTATACCAAGGCCCTTAAGTTATTGGAAG 5113
1473IleIleLeuAspThrLeuGlu..... 1479
5114 AATAAGAACTTCAGTGCACCTTCATATGCTGACCTGAAATTAACACA 5163
1480 IleValValGlnThrValSerValThrGluSerLysGluSerIleLeuG1 1496
5164 CTGGTAGAAGATTGTGTCAAAAGTTCGATCAGCAGCTGTTACCTGTTTGA 5213
1496 yGlyValLeuLysValLeu.....LeuHisSerM 1506
5214 AAACATTTTAGCCACAAGAACTGCACATAGTTTCGGAGAGATTTATAAGA 5263
1506 etAlaCysAsnGlnSerAlaValTyrLeuGlnHisCysPheAlaThrGln 1522
5264 TGACAACAGATCCAATGCTGGCCCTATCTACAGCCT...TTTAGAACAATCA 5310

1523 ArgAlaLeuValSerLysPheProGluLeuLeuPheGluGluGluThrG1 1539
5311 AGAAAAAGTTTTTTAGAGTACCAGATTGTGACAAAGAAAAACCTTTTGA 5360
1539 uGlnCysAlaAspLeuCysLeuArgLeuLeuArgHisCysSerSerSerI 1556
5361 AGGCGTGGATGATATA..... 5376
1556 leGlyThrIleArgSerHisProSerAlaSerLeuTyrLeuLeuMetArg 1572
5377AATCTGTGATTCCTTAAGT 5397
1573 GlnAsnPheGluIleGly.....AsnAsnPh 1581
5398 GAAATCATGACATTTTGATAAAGACACTGACTGTGCTTTTGTGGACAG 5447
1581 eAlaArgValLysMetGlnValProMetSerLeuSerSerLeuValGlyT 1598
5448 TGGAGGCACAAAATGTGAAATTTCTCAATTATTAAAGCCAATGTGTGAAG 5497
1598 hrSerGlnAsnPheAsnGlnGluPheLeuArgArgSerLeuLysThrIle 1614
5498 TGAAAACTGACTTT.....TGTCCAGACTGTGA 5523
1615 LeuThrTyrAlaGluGluAspLeuGluLeuArgGluThrThr..... 1628
5524 CTTCATACTTGTATTCATGATATTTTACTCCAGATACAAATGAATCATG 5573
1629PheProAspGlnValGlnAspLeuValPheAsnLeuHisM 1642
5574 GAGAAATCTGCTTTTCTACACATGTTCAGGATTTTTCACCAGCTGTCTTC 5623
1642 etIleLeuSerAspThrValLysMet..... 1650
5624 GACACTTTCGCAAAACAGCCGATCCACAACCCCTGCAAACTTGATTCA 5673
1651LysGluHisGlnGluAs 1656
5674 GAGTCAGACACTTTTTTCCGATGCTGTTGGATAAAAAATCACAAAGAAC 5723
1656 pProGluMetLeuIleAspLeuMetTyrArgIleAlaLysGlyTyrGlnT 1673
5724 AATGCTTGCTGTGTGACTACATGAGAAGA.....CAA 5758
1673 hrSerProGluArgLeuThr.....TrpLeuGlnAsn 1683
5759 AGAGACCTTCTTCAGAACAAATTTTAATGATGCTTTCGTGGCTGGAT... 5805
1684 MetAlaGlyLysHisSerGluArgSerAsnHisAlaGluAlaGln.. 1699
5806TTAAATTATCTAGAGTTGCCAAGGT 5831
1700CysLeuValHisSerAlaAlaLeuValAlaGluTyrLeuS 1713
5832 AGCTCAGTCTGTGCTGCTCCTTACAGCTTTACTC.....TATGCAG 5875
1713 erMetLeuGluAspArgLysTyrLeuPro..... 1722
5876 AAATCTATGCCAGATAAGAAAAAGTATGATGATCAAGAGAAAGAAGTCTT 5925
1723ValGlyCysValThrPheGlnAsnLleSerSerAs 1734
5926 GCATTTGAGAAGAAGACCCAGAGTACAACTATTTCTAGCTTGAGTAAAA 5975
1734 nValLeuGluGluSerAlaValSer..... 1742
5976 AAGTAAGAAAGAACTGGAATAGTTTACAGGATCTTCCTTAGAAATCT 6025
1743 .AspAspValValSerProAspGluGluGlyIleCysSerGly..... 1756
6026 ACAGAAGTATAGGGGAGCCAGATAGTTTGTATGCGCTGCTGGAGGGAAG 6075
1757LysTyrPheThrGluSerGlyLeuVa 1765

```
6076 ATGTACAACCCATTACTAGACTACGACATATGACACGACGACATGTG 6125
      :::::
1765 IGlYLeu.....LeuGlulnAla..AlaIaSerP 1775
      ||| ||||| ||| ::|||
6126 GGGCAAGCCCTAGTAACATATGACCTGCAACAGCAATCCCTCATCAA 6175
1775 heSerMetAlaGlyMetYrGluaIaValasn..... 1785
      ::|||::: ::|||:::
6176 CAGCCAGGCAAGATCATTCAGCCCTGCAGAAATTGGACCTGCCAT 6225
1786 .....GluValTyrLysValLeuIleProIleHisGluAlaAsnArgAs 1800
      ::|||::: ||| ::: |||::: |||
6226 ATTCTTCCGCTAT.....TTAAAGGATGGATTATGAATAAAGA 6269
1800 p..AlaLysLysLeuSerThrIleHisGlyLysLeuGlnGluAlaPhe. 1815
      |::: ::|||::: ::||| ||| |||:::
6270 CTGGTGCTCCTGACTAGAAGAACTTCAT.....TACCAAGCAGCATGA 6313
1816 .....SerLysIleValHisGln 1821
      ||||| |||:::
6314 GGAATATGACAGTGGGACCATTCACACTTCGCTCAGCAAGAAGTAGAAGA 6363
1822 SerThrGlyTyrPgluArgMetPheGlyThrTyr..... 1832
      ::::: ||| :::::
6364 ACCAGTTACCATGAATCATTTGTACAATGCTCTACAATCTCTAAGAGACAG 6413
1833 ...PheArgValGlyPheTyrGlyThrLysPheGlyAspLeuAspGluG 1848
      ||| ::: ::: |||::: ::: |||
6414 AGAATTCTCTACATTTTATGAAGTCTCAAAATATGCCAGATAAAGAAG 6463
1848 lngluPheValTyrLysGlu.....Pro 1855
      ||| ::: |||:::
6464 TGGAAAGAGATGCTAAGCGCAGCCCTGAGTCTGTATTCGCTCATGCC 6513
1856 AlaIleThrLysLeuAlaGluIleSerHisArgLeuGluGlyPheTyrGl 1872
      ::::: |||::: |||::: ::|||:::
6514 ACACTTAGCAGGTTGCCAGGCCATTGGA...GAGCTGGAAGCATT...GG 6557
1872 yGluArgPheGlyGluAspValVal..... 1880
      |||| |||::: |||:::
6558 GGAGCTTTTCTCAAGATCAGTCACACATAGACAACCTCTCTGAAGTATATA 6607
1881 .....GluValIleLysAspSer..... 1886
      ::::: |||||
6608 TTAAGTGGCAAAACACTCCAGCTTCTCAAGGACAGTGAATTTTACTTTT 6657
1886 ..... 1886
6658 CAGGAGCCTATCATGGCTCTACGACAGTCATTTTGAGATCCTGATGGA 6707
1887 .....AsnProValAspLysCysLysLeuAsp..ProAsnL 1898
      ||| ::||| ||| ::|||
6708 AAAGGAATGACAACTCACAAGAGATGTATTAAAGACATTCACCA 6757
1898 ysaIaTyrIleGlnIleThrTyrValGluProTyrPheAspThrTyrGlu 1914
      ||| ::|||::: ||| ::: |||:::
6758 AACACCTGTAGAACTCTATACTGGCCAGAACTTCAAGAAGAACTCAG 6807
1915 MetLysAspArgIleThrTyrPheAspLysAsnTyrAsn 1927
      ::||| ::: |||::: |||
6808 CTCCCTGAAAGGCAATATTTCAAAATTAACAGTACAAT 6846
seq_name: /cgn2_6/ptodata/2/lna/5B_COMB.seq:US-08-822-445-11
seq_documentation_block:
; Sequence 11, Application US/08822445
; Patent No. 5952223
; GENERAL INFORMATION:
; APPLICANT: Kaplan, Jerry
; APPLICANT: Perou, Charles
; APPLICANT: Moore, Karen
; TITLE OF INVENTION: COMPOSITIONS FOR THE DIAGNOSIS
```

```

; TITLE OF INVENTION: AND TREATMENT OF CHEDIAK-HIGASHI SYNDROME
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036/2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/822,445
; FILING DATE: 21-MAR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7853-062-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1225 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 190..11208
; US-08-822-445-11

alignment_scores:
      Quality: 150.00      Length: 2113
      Ratio: 0.177      Gaps: 101
Percent Similarity: 40.038      Percent Identity: 18.599

alignment_block:
US-09-737-246-2 x US-08-822-445-11 ..

Align seg 1/1 to: US-08-822-445-11 from: 1 to: 12225
56 AsproValAspleuGluAspTyrLeuIleThrHisProLeuAlaValAs 72
      ||| ||||| |||::: |||::: ::|||:::
3898 GATGGGGTAGACTTAAGTCTGAA.....ACAGAAAGTTTCAGTGCATC 3941
72 pSerGlyProLeuArgAspleuIleGluPheProProAspPileGluV 89
      ::|||::: |||::: |||::: |||:::
3942 AAGCAGTCCA..AATGACTTACTCGAATAAATTTGCTTCTGCT 3985
89 alValTyrSerProArgAspCysArg.....ThrLeuValSerAla 102
      ::|||::: |||::: |||::: |||:::
3986 TAAATTAT...CCTGAGATTTGTATGCTGGAATTAAATTTGCTTCTGCT 4032
103 ValProGluGluSerGluMetAspProHisValArgAspCysIleArgSe 119
      ::: ::|||::: |||||
4033 AGTAAAGCCAAACTGTGATGTGCTTGCCCATGTA..... 4065
119 rTyrThrGluAspTyrAlaIleValIleArgLysTyrHisLysLeuGlyT 136
      |||::: |||::: |||::: |||:::
4066 ....TTTGAGAGTTTTTGAAATATATTAGG..... 4092
136 hrcGlyPheAsnProAsnThrLeuAspLysGlnLysGluArg..... 149
      ||||| |||::: |||::: |||:::
4093 .....CAGAAAGAAAGAAATGTTT 4113
```


150GlnLysGlyLeuProLysGlnVal.....PheGluSe 160
 ||||:|||| ||||:|||| ||| ||
4114 CTGCTCATGCAACAGGAACTGTGAAAAATCTTTAGGAGGGTCTTGAG 4163
160 rAspGluAlaProAspGlyAsnSerTyrGlnAspAspGlnAspLeuL 177
 | :: ||| :::|||||
4164 TATTTTAAACACAGGATGATTTCTGATTTTCAA..... 4194
177 ysArgArgSerMetSerIleAspAspThrProArgGlySerTrpAlaCys 193
 |||||
4195GCATGC 4200
194 SerIlePheAspLeuLysAsnSerLeuProAspAlaLeuLeuProAsnLe 210
 :::||||| ||| ||| |||:|:| ::
4201CAGAGAGTATGGTGATCTTTGGTATCTTGAT 4235
210 uLeuAspArgThrProAsnGluGluIleAspArgGlnAsnAspAspGlnA 227
 :::||||| :::||||| ||| :::|
4236 GAGTTCAGAACAATGTTCAAGAAGCT.....AACCTTCCTTTG 4275
227 rGlySerAsnArgHisLysGluLeuPheAlaLeuHisProSerProAsp 243
 :::||||| :::||||| |||
4276 AGAATATTTCTGGAGAAATCTCCTGTGACAAAAATCTTCTTCTGGGTAT 4325
244 GluGlu.....GluProIleGluArgLeuSe 252
 :::||||| |||||:|:| ::
4326 TCTGAAAAATTATGAAAGTAGTACTACTATGAGCCCTTCACAGTATCTAA 4375
252 rValProAspIleProLysGluHisPheGlyGlnArgLeuValLysC 269
 :::||||| :::||||| ||| ||| |||
4376 CCTTCCCTTACTGACGCTCCAAATTTAAGCAACGGTGTTCATCACAA 4425
269 ysLeuSerLeuLysPheGlu.....IleGluIleGlu 279
 :::||||| |||:|:| ||| ||| |||
4426 AAGTATCCTGGGATTTTAAACAGTAAGCCCATGGCTTTATGAGAAGAGC 4475
280 ProIlePheAlaSerLeuAlaLeuTyrAspValLysGluLysLysIle 296
 |||
4476 ACGAGTTTC..... 4484
296 eSerGluAsnPheTyrPheAspLeuAsnSerGluGlnMetLysGlyLeuL 313
 :::||||| :::||||| |||
4485ACGGAGCAAGAAAGAGGCTGAT 4506
313 euArgProHisValProProAlaAlaIleThrThrLeuAlaArgSerAla 329
 ||| ||| |||:|:| ||| |||:|:|
4507 AGAGAGAGTTTCCCAATCGGCTGCTTC...ATCTTGGCACATAGC... 4550
330 IlePheSerIleThrTyrProSerGlnAspValPheLeuValIleLysIle 346
 ||||| :::|:| ||| |||
4551CCCAGTCCACCTGCCGTTGCTGGGCCAAACT 4582
346 uGluLysValLeuGlnGlnGly.....AspIle 355
 | :::||||:|:| :::|:| |||
4583 GCTGCCACACACCTATTCAGAAAGTTTCAAGTGTTCCTGTTAATGTG 4632
356 GlyGluCysAlaGluProTyrMetIlePheLysGluAlaAspAlaThrIly 372
 ||||| :::|||||:|:| |||
4633 ...GAGTGT.....ATCCATGAAGCTGAGAGTACTAC 4661
372 sAsnLysGluLysLeuLysLysLeuLysSerGlnAlaAspGlnPheCysG 389
 :::||||| ||| :::|:| |||
4662 AGAAAAAGGAAAGAAAGTAAAGAAAAAGAAACAATCA..... 4698
389 InArgLeuGlyLysTyrArgMetProPheAlaTrpThrAlaIleHisLeu 405
 |||
4699TTA 4701
406 MetAsnIleValSerSerAlaGlySerLeuGluArgAspSerThrGlu.. 421
 :::||||| :::||||| ||| |||
4702 ATTTTACCAGATAGCAGTTTGATGTGTACAGAGAGACAGACAGACAAGG 4751
422 ...ValGluIleSerThrGlyGluArg..... 429

4752 TGCAGAGTACATAAATCTGTGTGAAGACTCATAGAGAAGAGATGATTC 4801
 ||||:| |||||
430LysGlySer.....TrpSerGlu 435
 ||||| |||:|:| |||:|:|
4802 ATATTAATTTCACTGGGATCCAAAGCGTTGATGATCCAAAGTGTGGCTGAT 4851
436 ArgArgAsnSerSerIleValGlyArgArgSerLeuGluArgThrThrse 452
 :::||||:|:| ||| :::|:| |||
4852 CCCCAACAATGCCACCTCTTATCTTTCGTGTGTCATGCAAT.....TC 4892
452 rGlyAspAsp.....AlaCysAsnLeuThrSer..... 461
 |:|:|:|:| ||| |||:|:| |||
4893 AAATGATGACATGAAGCTGTTTCTAGCACAGCTTGAATCACAGAGAG 4942
462PheArgProAlaThr.....LeuThrValThrAsnPhePhe 473
 ||| |||:|:| |||:|:| |||
4943 ATATTTTCTCCCAAGCAAAATGGCAACATTTAGTACTCACCTACTTACAG 4992
474 LysGlnGluGlyAspArgLeuSerAspGluAspLeuTyrLysPheLeuAl 490
 :::|:|:| ||| :::|:| :::|:|
4993 CAGCCCAAGGAAAGAGAGATTCATGGGAAAAATCTCATATAGGTCCTC 5042
490 aAspMetArgArgProSerSerValLeuArgArgLeuArgProIleThra 507
 :::|:|:|:|:|:| :::|:| |||
5043 TGGACAGAGGAAGCCTGATGTTACTTTGGATTTTATGCTTCCAGAAAAA 5092
507 laGlnLeuLysIleAspIleSerProAlaProGluAsnProHisTyrCys 523
 :::|:|:|:|:| ||| :::|:| |||
5093 CAAGTTTGTCACTGTGAT.....AGCAATTAACAATTTTGC 5127
524 LeuThrProGluLeuLeuGlnValLysLeuTyrProAspSerArgValAr 540
 :::|:| |||:|:| :::|:|
5128 ATGATTTGCCATTTGTTTATCAATCCCAA..... 5154
540 gProThrArgGluIleLeuGluPheProAlaArgAspValTyrValProA 557
 :::|:|:|:|:|:|:|:| :::|:| |||
5155GAAGAGTTTGTGACGTTGGCTGAAAAA..... 5181
557 snThrThrTyrArgAsnLeuLeuTyrIleTyrProGlnSerLeuAsnPhe 573
 ||||| |||
5182 ..TGGGACCTGGGAATTTGCTTCTCTTC..... 5208
574 AlaAsnArgGlnGlySerAlaArgAsnIleThrValLysValGlnPheMe 590
 :::|:|:|:|:|:| :::|:| |||
5209AACGGAGCTTAAGTTGTTGTTCAACAAGAGCCTTTTATCT 5246
590 tTyr.....GlyGluAspProSerAsnAlaMetProValIlePheGlyL 605
 ||| ||| ||| :::|:| ||| |||
5247 GTATGCTTGTGGACCCCAACCATATCATCTGTAAATGCCATGTAAGTATGCA 5296
605 ysSerSerCysSerGluPheSerLys..... 613
 || :::|:|:|:|:| |||
5297 AG...CCAGTCAATGACTACTCCAAATATATTAATAAGAAATTTTGCA 5343
614GluAlaTyr..... 616
5344 TGTGAACAATAACAGAGACTTTTATGACCAAGAAGATGTGATATTGG 5393
617ThrAlaValValTyrHisAsnArgSerProAsp. 627
5394 TCTCTTAATTGAAGTCTTTCAAGTTGTTTATACAACTTACTGTCTGCTC 5443
628PheHisGluGluIleLysValLys..... 635
 :::|:|:| |||:|:| |||
5444 AGTATACATCTATGAACACAGTGAATTAGACTTAAAGGTCAAAATGAANAAC 5493
636LeuProAlaThrLeuTh 641
5494 CAACCTCTCTCAAAAGACCCTTACGCTCAAAAGAGTTTCAGACATCTTATT 5543
641 rAspHisHisHisLeu..... 646
 :::|||||

5544 AGAACCTCATCATCTAAAAAATCTCCAACTACTGAATATAAACTATTC 5593
647LeuphethrPhe 650
5594 AAGCATCTTCGACGAAATTTGGTGAAGTGGCATATTGTTTCTCTTT 5643
651 TyrHisVal.....SerCysGlnGln..... 659
5644 GCCAGGTTGTGAAGTCAAGCACTCAAGCATTTAGCACT 5693
659 659
5694 GCGAGTTACTCTCATTAATTAACAACCAAGAGTACATGAAT 5743
659 659
5744 TAGAAAAATTGTAATGACTTTCTATGATTCATCAGGTGTTGATCAAAACA 5793
660AsnThrProLeuGluThrProva 667
5794 AAATGCATTTGTTGGTTTACATTTTGAAGACCTCTTGAAGAGTGCCTG 5843
667 IGlyTyrThrTrpIlePrometLeuGlnAsnGlyArgLeuLysThrGlyG 684
5844 TGGTGAAGATATTATTATGAAAGAGAAAT.....GGAG 5878
684 LnPheCysLeuProValSerLeuGluLysProProGlnAlaTyrSerVal 700
5879 AGTTTAAGTTGGATGTAGACTCTAATGCTATAATCCAAAGATGTTAAGCTG 5928
701 LeuSerProGluValProLeuProGlyMetLysTrpValAspAsnHisLys 717
5929 TTAGAG...GAACATAATGCTTGCAGTGAAGATATGAGTAAAGCAGAGCA 5975
717 sGlyValPheAsnValGluValAlaValAlaValSerSerIleHisThrGlnA 734
5976 AGGTGTTGGGAAACTTTGCTAGACGCTTAGAAGTCCCTCATCAGAGCAG 6025
734 sPProTyrLeuAspLysPhePheAlaLeuValAsnAlaLeuAspGluHis 750
6026 ATCACCACGACGAGATGTT..... 6045
751 LeuPheProValArgIleGlyAspMetArgIleMetGluAsnLeuGlu 767
6046AATATTAA 6053
767 uAsnGluLeuLysSerSerIle..... 774
6054 GCAGTTATGAAAGCTCAAGTGGTTCATCACTTCTACTGACTTGTGAGG 6103
775SerAlaLeuAsnSerSerGlnLeuGluProValValArg..... 787
6104 TTTTGCAGGAATACAAAGAGGGGCACTCACACCCATGCCCGCAGAGGTT 6153
788PheLeuHisLeuLeuLeuAspLysLeu..... 796
6154 TGTAGATCATTTGTGAAAATTATAGCAGAAGTCCCTGGATCTCCTCCAGA 6203
797IleLeuValIleArgProp 804
6204 TTTGGAATTATTGACAATTATCTCAATTTCTTTAGCAGTTCAACCTC 6253
804 roVal..... 805
6254 CTACTAATACTTACGTTTGCACAATCCACGAACTTCTACTTTCTTTG 6303
806 ..IleAlaGlyGlnIleValAsnLeuGlyGlnAlaSerPheGluAlaMe 821
6304 CACATAGATGGCAAGATCTTTCAG.....GAGAAAGT 6335
821 tAlaSerIleIleAsnArgLeuHisLysAsnLeuGluGlyAsnHisAspG 838
6336 GCGGTCAATCATGTACCTG.....AGGCATTCCA 6364

838 LnhISGlyArgAsnSerLeuLeuAlaSerTyrIleHisTyrValPheArg 854
6365 GCAGTGGAGGAGGTCCTTATGAGCCCT....GGATTATAGTAATA 6408
855 LeuProAsnThrTyrProAsnSerSerProGlyProGlyGlyLeuGlu 871
6409 AGCCCATCTGTTT....ACTGCTTCACCATATGAAGAGAGAAATTC 6452
871 yGlySerValHisTyrAlaThrMetAlaArgSerAlaValArgProAlaS 888
6453 CTCTAATATATTCCACAACAGATGGCCGCATATGCTGCGTCTAGAA 6502
888 erLeuAsnLeuAsnArgSerArgSerLeuSerAsnSerAsnProAspIle 904
6503 GCCTACACGATTCCTTACTTCTTCACTACTAACGCAATCACAAAACTG 6552
905 SerGlyThrProThrSerProAspAspGluValArgSerIleIleGlySe 921
6553 ACTGGAAGTTGGGTTGTAGTATCGACAGGTTACAAATATT..... 6594
921 rLysGlyLeuAspArgSerAsnSerTrpValAsnThrGlyGlyProLysA 938
6595GCAGATACTTATGTTGCCACCACCAATCAAGAAGAAC 6628
938 laAlaProTrpGlySerAsnProSerProSerAlaGluSerThrGlnAla 954
6629 AAAATTCCTTGGGAGTTCGCACACACTGAAAAAAGCAAGAGAGACGCA 6678
955 MetAspArgSerCysAsnArg..... 961
6679 TTCATCAGTAGCTGTGAGTCTGCAAAACTGTTTGTGAATGGAAGCTGT 6728
962 .MetSerSerHisThrGluThrSerSerPheLeuGlnThrLeuThrGlyA 978
6729 CCTCTACGCCAGTCTCTGTCAAGTATGCCCAAGGAGTGCCTGGGA. 6777
978 rGLeuProThrLysLysLeuPheHisGluGluLeuAlaLeuGlnTrpVal 994
6778 ..TTCCAGTGGTCAAGCAGATCATTAACAGTTGGGAGCAGAA..... 6819
995 ValCysSerGlySerValArgGluSerAlaLeuGlnGlnAlaTrpPhePh 1011
6819 6819
1011 ePheGluLeuMetValLysSerMetValHisHisLeuTyrPheAsnAspL 1028
6820CCAGGTCAG 6829
1028 ySLeuGluAlaProArgLysSerArgPheProGluArg.....PheMet 1042
6830 AAGATGACAGTCTGGGATGAGTCCCGCCACGCGCAGCTGATTACCTA 6879
1043 AspAspIleAlaAlaLeuVal.....SerThrIleAlaSer..... 1054
6880 AAGGATTTGGCTCTCTTCCAGCGAAGCCACAGCACTATTGCCAACCCTGG 6929
1055A 1055
6930 GCTAGCTTTTCTTCCACAGAACGAGTGTGAGCTGTGGCCGTGGCCAA 6979
1055 spIleValSerArgPheGlnLysAspThrGluMetValGluArgLeuAsn 1071
6980 GTCTTGTGTAGAAACACTGATGATTCGGA..... 7011
1072 ThrSerLeuAlaPhePheLeuAsnAspLeuLeuSerValMetAspArgGl 1088
7012 ..AACTTGGCTATTCTCTTGGTTATGAGCCAAATTACAAACGAACTGC 7058
1088 yPheValPheSerLeuIleLysSerCysTyrLysGlnValSerSerLysL 1105
7059 AAGTGCTCACAGTGTAACTGAAGACTGTGTTGGTACCTATATGCTGTGAT 7108

1105 euTyrSerLeuProAsn.....ProSerValLeuVal 1115
||||:||||| :||:|||||:|
7109 TATATGAACTCCTAAGTGGGTTCTTCTTATCCTGCCGTGATGTTTGCTT 7158
1116 Ser.....LeuArgLeuAspPheLeuArgIleIleCy 1126
:|:|:| :|:|:| :|:| :|:|:|
7159 GAAGATGTGATGGACAAGCTTATTCAAGCAGATACACTTTTGCTCCTGT 7208
1126 sSerHisGluHisTyrValThrLeuAsnLeuProCysSerLeuLeuThrP 1143
:|:|:| :|:|:| :|:|:| :|:|:|
7209 TAACCAC..... 7215
1143 roProAlaSerProSerProSerValSer..... 1152
|||||:|:|:|:|:|:|:|:|
7216CCATCACCACCTATACACAACAAGGTGTTAATACTATTA 7254
1153SerAlaThrSerGlnSerSerGlyPheSerTh 1163
:|:|:|:|:|:|:|:|:|:|:|:|
7255 GATGCATATTTTGTAGAGCATCTAAGACAAAAAGATAATTCTGAA 7304
1163 rAsnValGlnAspGlnLysIleAlaAsnMetPhe..... 1174
||| :|:| :|:|:|:|:|:|:|:|
7305 GAATCGTGGATTTTCCCTTGCTAGCCAAACAGTTGTATCTTCATCGAGGAA 7354
1175GluLeuSerValProPheArgGlnGlnHisTyrLeuAlaGlyLeu 1189
||||| :|:|:|:|:|:|:|:|:|:|
7355 CTCAGAATGTTGTAGAATGCTTTCATCGAAATGTTCTTGTCGACATATT 7404
1190 ValLeuThrGluLeuAlaValIleLeuAspProAspAlaGluGlyLeuPh 1206
||| ||| :|:| :|:| :|:|:|:|:|
7405 GGCCTTGATGAAGAATTGTGATCTGGAAGATGTGAGAAACATGGATTGTT 7454
1206 eGlyLeuHisLysLysValIleAsnMetValHisAsnLeuLeuSerSerH 1223
| :|:|:|:|:|:|:|:|:|:|:|:|
7455 T..... 7455
1223 iAspSerAspProArgTyrSerAspProGlnIleLysAlaArgValAla 1239
:|:|:|:|:|:|:|:|:|:|:|:|
7456CAGAAGTGGTCT..... 7467
1240 MetLeuTyrLeuProLeuIleGlyIleIleMetGluThrValProGlnLe 1256
:|:|:|:|:|:|:|:|:|:|:|:|
7468GTCATTCCCTATTCTGGGACTAATA..GAGACC.....TCTCT 7502
1256 uTyrAspPheThrGluThrHisAsn..... 1264
||||| :|:|:|:|:|:|:|:|:|:|
7503 AATGACAACATACTCTTCGATAATGCTCTTTACTTCTTCCCACATCATG 7552
1265GlnArgGlyArgProIleCysIleAlaThrAspAspTyrGlu 1278
||:|:|:|:|:|:|:|:|:|:|:|:|
7553 CAGTAGTTCAAAAGCGGAAA..... 7572
1279 SerGlnSerGlySerMetIleSerGlnThrValAlaMetAlaIleAlaG1 1295
:|:|:|:|:|:|:|:|:|:|:|:|
7573AGCATTTGCTGG 7583
1295 yThrSerValProGlnLeuThrArgProGlySerPheLeuLeuThrSerT 1312
| :|:|:|:|:|:|:|:|:|:|:|:|
7584 TCCTCGAAATTTCCCTTCCTGCTCAAACTGAATCCCTTCTGATGAAAAATGC 7633
1312 hrSer.....GlyArgGlnHisThrThrPheSerAlaGluSerSer 1325
||| :|:|:|:|:|:|:|:|:|:|:|:|
7634 GTTCAGTGGCAAAATGATGAGCTTCATGTGATGATGCAACGAGAAATGAGC 7683
1326 ArgSerLeuLeuIleCysLeuLeuTrpValLeuLysAsnAlaAspGluTh 1342
:|:|:|:|:|:|:|:|:|:|:|:|
7684 CAAGAG.....AACCTAGCCCAAGCAACTGAAC 7712
1342 rValLeuGlnLysTrpPheThrAspLeuSerValLeuGlnLeuAsnArgL 1359
| ||| :|:|:|:|:|:|:|:|:|:|:|:|
7713 GGAACCTTGGCGAGAGACTACAGAGGCTCACCTGTTTTAGCAGTCAACAGGA 7762
1359 euLeu.....AspLeuLeuTyrLeuCys..... 1366

7763 TTATTTATCAGAATTTTAATTCAGACATTTATTGACATTTTGAGACTCCA 7812
:|:|:|:|:|:|:|:|:|:|:|:|
1367ValSerCysPhe.....GluTyrLysG1 1374
:|:|:|:|:|:|:|:|:|:|:|:|
7813 GAAATGTACTCAAGACCAAGACCCTCAGTTTTCAGACCGAAATTTCTGA 7862
1374 yLysLysValPheGluArgMetAsnSerLeuThrPheLysSerLysA 1391
:|:|:|:|:|:|:|:|:|:|:|:|
7863 GGAATAATATCATCATGAAACAGCTTCTGTGTTTCAATCCATTCAGAAAG 7912
1391 sPmetArgAlaLysLeuGluGluAlaIleLeuGlySerIleGlyAlaArg 1407
:|:|:|:|:|:|:|:|:|:|:|:|
7913 AAATTTTACATATCTGTGAGAGGATTCAAAGTATCTATTGCT..... 7956
1408 GlnGluMetValArgArgSerArgGlyGlnLeuGluArgSerProSerG1 1424
|||:|:|:|:|:|:|:|:|:|
7957TCAAG 7961
1424 ySerAlaPheGlySerGlnGluAsnLeuArgTrpArgLysAspMetThrH 1441
:|:|:|:|:|:|:|:|:|:|:|:|
7962 TAAAGCCAGTGGTTCCAAGCAG.....CAATGCACTAAATTTCTG... 8001
1441 iStrParG..GlnAsnThrGluLysLeuAspLysSerArgAlaGluIle 1456
||| :|:|:|:|:|:|:|:|:|:|:|:|
8002 ..TGGTCTTGTAAGGAGACCTTCGGAATGCAGCTTGAGGAGCTACTAGTG 8049
1457 GluHisGluAlaLeuIleAspGlyAsnLeuAlaThrGluAlaAsnLeuI1 1473
:|:|:|:|:|:|:|:|:|:|:|:|
8050CATATTTTGTGCGCCAGCCCGCTGCACACAAGAGAGAAACAAAT 8093
1473 eIleLeuAspThrLeuGluIleValIleGlnThrValSerValThrGluS 1490
| :|:|:|:|:|:|:|:|:|:|:|:|
8094 T.....TTGAAATAGTTCATGAA.....CCAA 8116
1490 eRLysGluSerIleLeuGlyGlyValLeuLysValLeuLeuHisSerMet 1506
:|:|:|:|:|:|:|:|:|:|:|:|
8117 ATCATCAGAAATACTACGAGACTGTCTCAGCCCATCCCTACACATGGA 8166
1507 AlaCysAsnGlnSerAlaValTyrLeuGlnHisCysPheAlaThrGlnAr 1523
||| :|:|:|:|:|:|:|:|:|:|:|:|
8167 GCC.....AAGTTAGTTTGTGATTTGTCAAG..... 8193
1523 gAlaLeuValSerLysPheProGluLeuLeuPheGluGluGluThrGluG 1540
||:|:|:|:|:|:|:|:|:|:|:|:|
8194 ...TTGATACATAATCACCACAGGTGAATTGACTGAAGAGAGACTA...G 8236
1540 lncYsAlaAspLeuCysLeuArgLeuLeuArgHisCysSerSerSerIle 1556
||:|:|:|:|:|:|:|:|:|:|:|:|
8237 GCACAGCAGAACTGCTTATGAATGCTTTGAAGTTATGTGTCAACAAGTGC 8286
1557 GlyThrIleArgSerHisProSerAlaSerLeuTyrLeuLeuMetArgG1 1573
:|:|:|:|:|:|:|:|:|:|:|:|
8287 ATCCCTGCCAGTGCATCAACAACCAAGCCTTATTAATAATGATCAAGA 8336
1573 n.....AsnPheGluIleGlyAsnAsnPheAlaArgValL 1585
:|:|:|:|:|:|:|:|:|:|:|:|
8337 GGAACAACAAAGAAATATGAAACTGAAGAAGAGTGAATTAAGCTGCTTGGC 8386
1585 ysmetGlnValPrometSerLeuSerSerLeuValGlyThrSerGlnAsn 1601
:|:|:|:|:|:|:|:|:|:|:|:|
8387 AGAAAACAGTT.....AAC 8400
1602 pheAsnGluGluPheLeuArgSerLeu.....LysThrI1 1614
||:|:|:|:|:|:|:|:|:|:|:|:|
8401 AATAATCAACAAAGTCTCTTTCAGCGCTGTGATTCAAAATCAAAAGATAT 8450
1614 eLeuThrTyrAlaGluGluAspLeuGluLeuArgGluThrThrPheProA 1631
| ||| :|:|:|:|:|:|:|:|:|:|:|:|
8451 ATCTAAATAGCTGCAGATATATCACCCAGCAGTGTCTCTCCCAAGGAA 8500
1631 spGlnValGlnAspLeuValPheAsnLeuHisMetIleLeuSerAspThr 1647
:|:|:|:|:|:|:|:|:|:|:|:|

8501 ATGAGAGAAAAAGGTGATCCAGCATATTAGAGGAATGTATAAAGTAGAT 8550
1648 ValLysMetLysGluHisGlnGluAspProGluMetLeuIleAspLeuMe 1664
::: ::::: ||| ::::: ::::: |||
8551 TTGAGTGCACGACGACACTTGGCAG....GAACCTATTCAGCAGCTGAC 8594
1664 tTyrArgIleAlaLysGlyTyrGlnThrSerProGluArgLeuThrTrp. 1680
::: ||| ::::: ::::: |||
8595 ACATGATAGACGACGATGTGTATGACCCCATCTACTATCCAACCTCATGGC 8644
1681 ..LeuGlnAsnMetAlaGlyLysHisSerGluArgSerAsnHisAlaGlu 1696
||| ::::: ||| ::::: |||
8645 AGTTGGATCCAAACAGAGGCCAAATCGAGAGAG....AGACGT 8685
1697 AlaAlaGlnCysLeuVal..... 1702
::: ||| :::
8686 TTACAGAGATGTTATTAACTATTCCAATAAGTAATCTCCTTAGGATAG 8735
1703HisSerAlaAlaLeuValAlaGluTyrLeuSerMetLeuGluAspa 1718
||| ::::: ||| ||||| ||| :::
8736 ACAGAAATCAGAGAGATGTGTCAAAACCAACCACTCTTACCTGTTGAAG 8785
1718 rgLysTyrLeuProValGlyCysValThrPheGlnAsnIleSerSerAsn 1734
||| ||| ::::: ||||| :::
8786 ACAA.....ACTCATCTCTTCTTCTCTTCTACT 8814
1735 ValLeuGluGluSerAlaValSerAspVal..... 1745
||| ::::: ||| ::::: |||
8815 GTC...AAAGACAAGCTGCAAGTGAATCTATAGAGTGAATCGAAGATG 8861
1746ValSerProAspGluGlu.....GlyI 1753
||| ::::: ||| ||||| :::
8862 CATCAGTGTTCACCATCTAGAGACACAGCTGTGTAATTGTTACTAGTA 8911
1753 IeCysSerGlyLysTyrPheThrGlu..... 1761
||| ||| ||||| :::
8912 AATGT...GGAATGTATTGTGGAAGATAATGCTTGATACAGTTGAA 8958
1762 ...SerGlyLeuValGlyLeuLeuGluGlnAlaAlaAlaSerPheSerme 1777
||| ::::: ||| ||||| ||| :::
8959 AGTCGAGCCTTCAGGAGAGTTGGAACACAGCATCATTTCTTGACA.. 9006
1777 tAlaGlyMetTyrGluAlaValAsnGluValTyrLys 1789
||||| ::::: ||||| :::
9007TATGAAGAAATTAAAGAAGTTCACAAG 9033